

NÁVRH A TVORBA PREDIKTIVNÍHO MODELU PRO HODNOCENÍ MÍRY RIZIKA LÉČBY U CHRONICKÉ LYMFOCYTÁRNÍ LEUKÉMIE

DESIGN AND CREATION OF A PREDICTIVE MODEL FOR EVALUATING THE
RISK OF TREATMENT IN CHRONIC LYMPHOCYTIC LEUKEMIA

Bc. Zdeňka Krchová

DIPLOMOVÁ PRÁCE

Vedoucí diplomové práce: Ing Lukáš Peter

Ostrava, 2021

Abstrakt

Tato diplomová práce se zabývá návrhem prediktivního modelu pro hodnocení míry rizika léčby u chronické lymfocytární leukémie. V současné době je přes 75 % nemocných diagnostikováno v časném stádiu (Binet A, Rai 0) a tedy nepotřebují léčbu, pouze jsou sledováni. Nicméně cca polovina z těchto nemocných se postupně zhoršuje a léčbu v budoucnu potřebuje. Průběh nemoci u pacientů je individuální a jejich pravděpodobnost nutnosti léčby je jen stěží předvídatelný. Tvorba prognostického modelu do nasazení první léčby u těchto pacientů, u kterých je vysoké riziko progresu onemocnění, by bylo velmi prospěšné. V teoretické části se věnuji rozšíření AIPS-E (alternativní mezinárodní prognostický skórovací systém) o další faktory, jako je věk nad 65 let, komplexní karyotyp (CK), delecii 6q21, trisomii 12, *IgH* a delecii 13q14, které by mohly zpřesnit prognostické skóre do TTFT pro rané fáze CLL. Pro statistické zpracování a pro návrh modelu byla poskytnutá data pacientů z FN Olomouc, u kterých byla diagnostikována CLL. Pro praktické využití byl navržen prediktivní model pro hodnocení míry rizika léčby u chronické lymfocytární leukémie.

Klíčová slova: CLL, raná fáze, nové faktory, fuzzy logika

Abstract

This diploma thesis deals with the design of a predictive model for assessing the degree of risk of treatment in chronic lymphocytic leukemia. Currently, over 75% of patients are diagnosed at an early stage (Binet A, Rai 0) and therefore do not need treatment, they are only monitored. However, about half of these patients are gradually deteriorating and need treatment in the future. The course of the disease in patients is individual and their probability of needing treatment is hardly predictable. Creating a prognostic model for the first treatment in these patients, who are at high risk of disease progression, would be very beneficial. In the theoretical part I deal with the extension of AIPS-E (alternative international prognostic scoring system) by other factors, such as age over 65 years, complex karyotype (CK), deletion of 6q21, trisomy 12, *IgH* and deletion of 13q14, which could refine prognostic scores to TTFT for early phase CLL. Data for patients from the Olomouc University Hospital in whom CLL was diagnosed were provided for statistical processing and for model design. For practical use, a predictive model was proposed to assess the degree of risk of treatment in chronic lymphocytic leukemia.

Keywords: CLL, early phase, new factors, fuzzy logic

Tímto bych chtěla poděkovat Ing. Lukáši Peterovi, Ph.D., za odborné rady a vedení mé diplomové práce. Dále děkuji pracovnícům Hemato-onkologické kliniky LF a FN Olomouc Mgr. Heleně Urbánkové, Ph.D. a Mgr. Nele Kasalové za odborné konzultace a poskytnutí dat k vytvoření mé práce a Mgr. Janě Zapletalové, Dr. za odbornou konzultaci ve statistickém zpracování.

Obsah

Seznam použitých symbolů a zkratk	5
Seznam obrázků.....	6
Seznam tabulek.....	7
Úvod.....	8
1. Leukémie	9
1.1 Základní informace o krvi a kostní dřeni	9
1.2 Typy leukémií	9
1.3 Lymfatický systém a jeho funkce	10
1.4 Chronická lymfocytární leukémie (CLL).....	11
1.5 Příznaky	12
2. Vyšetřovací metody.....	13
2.1 Základní laboratorní vyšetření pro stanovení diagnózy.....	13
2.2 Zobrazovací vyšetření při stanovení diagnózy	15
2.3 Prognostická vyšetření z periferní krve.....	15
2.4 Klasifikace.....	21
2.5 Prognostické modely	22
2.6 Léčba	24
3. Návrh a tvorba prediktivního modelu pro hodnocení míry rizika léčby u chronické lymfocytární leukémie.....	25
3.1 Cíl práce.....	25
3.2 Soubor dat, statistika a metodika využitá k tvorbě prediktivního modelu	25
3.2.1 Datová sada.....	25
3.2.2 Statistické zpracování dat	26
3.2.3 Metodika a tvorba prediktivního modelu	41
Závěr.....	48
Seznam literatury a zdroje obrázků	49

Seznam použitých symbolů a zkratek

CLL	Chronická lymfocytární leukémie
ALL	Akutní lymfoblastová leukémie
AML	Akutní myeloidní leukémie)
CLL	Chronická lymfocytární leukémie
CML	Chronická myeloidní leukémie
AIPS-E	Alternativní mezinárodní prognostické skóre
BCR signalizace	B buněčný receptor
Proteinu Bcl-2	B-buněčný lymfom
DNA	Deoxyribonukleová kyseliny
<i>IgVH</i>	Těžký řetězec imunoglobulínu
FISH	Fluorescenční in situ hybridizace
IPI	Mezinárodní prognostický index
IPS-E	Mezinárodní prognostické skóre pro asymptomatické rané fáze chronické lymfocytární leukémie
AIPS-E	Externí validace mezinárodního prognostického skóre pro asymptomatická raná fáze chronické lymfocytární leukémie a návrh alternativního skóre
CK	Komplexní karyotyp
TTFT	Čas do první léčby
OS	Celkové přežití
Anémie	Chudokrevnost
Trombocytopenie	Snížení krevních destiček
Splenomegalie	Zvětšení sleziny
Lymfadenopatie	Zvětšení lymfatických uzlin
Lymfocytóza	Zvýšený počet lymfocytů
BM	Biologický materiál
KD	Kostní dřeň
PK	Periferní krev

Seznam obrázků

1.1 Oblasti mízních (lymfatických) uzlin v lidském těle.....	10
1.2 Lymfocyty u chronické lymfocytární leukémie	11
1.3 Výskyt chronické lymfocytární leukémie v populaci ve vztahu k věku	11
2.1 Chromozomy v buňkách chronické lymfocytární leukémie zobrazené při cytogenetickém vyšetření	13
2.2 Výsledek molekulárně genetické analýzy	14
2.3 Odběr kostní dřeně sternální punkcí.....	14
2.4 Odběr kostní dřeně trepanobiopsíí kosti kyčelní	14
2.5 Výsledek FISH vyšetření interfázních jader u nemocného s chronickou lymfocytární leukémií.....	17
3.1 Kaplan – Meierova křivka doby do začátku léčby u AIPS-E.....	35
3.2 Kaplan – Meierova křivka doby do začátku léčby u mod. AIPS-E + 6q21	40
3.3 Vyhodnocovací pravidla u fuzzy logiky	42
3.4 Fuzzy model základní rozdělení.....	45
3.5 Fuzzy model nastavení členství.....	45
3.6 Fuzzy model nastavení pravidel	46
3.7 Fuzzy model – prediktivní model	46
3.8 Fuzzy model rozložení plochy TTFT do 1 roku.....	47
3.9 Fuzzy model rozložení plochy TTFT do 5 let.....	47

Seznam tabulek

2.1 Klasifikace CLL podle Jacquese – Louise Bineta.....	21
2.2 Mezinárodní prognostický index	22
2.3 Mezinárodní prognostické skóre pro asymptomatické rané fáze chronické lymfocytární leukémie	23
2.4 Externí validace mezinárodního prognostického skóre pro asymptomatická raná fáze chronické lymfocytární leukémie a návrh alternativního skóre	23
3.1 Určení statisticky významných faktorů, které ovlivňují TTFT	32
3.2 Kategorizace dle AIPS-E	32
3.3 Kategorizace rizika podle AIPS-E skóre	32
3.4 C-statistika – AUC u AIPS-E.....	33
3.5 Výpočet 1 a 5letého TTFT u AIPS-E.....	33
3.6 Life Tables u AIPS-E.....	33
3.7 Odhad mediánu a průměrné doby do začátku léčby u AIPS-E	34
3.8 Kategorizace podle míry rizika u AIPS-E.....	34
3.9 Log Rank test u prognostického skóre AIPS-E	34
3.10 Určení statisticky významných nových faktorů, které ovlivňují TTFT	36
3.11 Parametry pro výpočet modifikovaného skóre AIPS-E	37
3.12 Kategorizace rizika podle modifikovaného skóre AIPS-E	37
3.13 Distribuce pacientů podle rizika	37
3.14 C-statistika – AUC u modif. AIPS-E.....	38
3.15 Výpočet 1 a 5letého TTFT u modif. AIPS-E + 6q21	38
3.16 Kategorizace rizika podle bodovacího skóre	38
3.17 Life Tables u modif. AIPS-E + 6q21	39
3.18 Odhad mediánu a průměrné doby do začátku léčby u modif. AIPS-E + 6q21.....	39
3.19 Kategorizace podle míry rizika u mod. AIPS-E + 6q21	39
3.20 Long Rank test u mod. prognostického skóre AIPS-E + 6q21.....	40
3.21 Určení hranic ACL podle Youdenovi metody	42
3.22 Určení hranic delece 11q a delece 17p podle Youdenovy metody	43
3.23 Kategorizace nízká, střední a vysoká hodnota	43
3.24 Kategorizace četnosti podle rozdělení nízká, střední a vysoká hodnota	43
3.25 Kontingenční tabulka rozdělení četnosti do rizikových skupin	43
3.26 Bodový odhad podílu pacientů s léčbou započatou do 1 – 5 let.....	44
3.27 Ověření prognostického modelu na datech AIPS-E + 6q21	44
3.28 Ověření přesnosti predikce na datech AIPS-E bez 6q21.....	44
3.29 Výpočet predikce v % pro jednoho pacienta	44

Úvod

Chronická lymfocytární leukémie (CLL) je nízce maligní lymfoproliferativní onemocnění postihující dospělé na západní polokouli. Jde o nejčastější leukémii v západním světě, kde tvoří téměř 30 % všech leukémií. Jedná se o hromadění lymfocytů v lymfatických orgánech (místních uzlin, slezina), v krvi a kostní dřeni. Zpravidla se nemoc vyskytuje u starších lidí, průměrný věk v době diagnózy se pohybuje mezi 65 a 72 lety. CLL se nejčastěji zjistí náhodně, a to u vyšetření krevního obrazu nebo u předoperačního vyšetření. Různorodý klinický průběh je pro CLL typickým rysem. Spousta nemocných nemusí být vůbec léčena, jsou pouze sledováni a dochází na pravidelné kontroly. Asi u poloviny nemocných dojde postupně k progresi nemoci a je nutné léčbu zahájit.

Tvorba co nejpřesnějšího prediktivního modelu, kdy léčbu zahájit, by byla velmi prospěšná. Neustále probíhají studie nových prognostických faktorů, které by napomohly určit, jakým směrem se nemoc bude ubírat. Zda povede či nepovede k progresi. V případě úspěšné tvorby prediktivního modelu do nasazení první léčby by vedlo k časnému nasazení léčby a tím prodloužení života pacienta.

Cílem této práce je vytvořit prognostický model, který zpřesní již navržený AIPS-E, kdy bude léčba zahájena. Budu vycházet z prognostického skóre AIPS-E a rozšířím toto skóre o prognostické faktory, které budou statisticky významné, jako je věk nad 65 let, komplexní karyotyp (CK), delecí 6q21, trisomii 12, *IgH* a delecí 13q14.

1. Leukémie

Leukémie (rakovina krve) je obecný název pro onemocnění kostní dřeně. Dochází zde k nekontrolovatelnému množení nádorových buněk, které neplní funkci zdravých buněk a nepodléhají zániku. Tím dochází k „utlačování“ krvetvorby ostatních buněk. Nádorové buňky se vyplavují do krve. Důvodem této nemoci jsou mutace (změny) genů v krvetvorných buňkách.

1.1 Základní informace o krvi a kostní dřeni

Krev se skládá z krevních buněk - červené krvinky, bílé krvinky, krevní destičky a plazmy.

Plazma tvoří tekuté médium pro krevní buňky. Je žluté barvy a má na starost tvorbu krevní sraženiny při poranění, zabraňuje infekci.

Červené krvinky (erytrocyty) patří mezi malé buňky, obsahují hemoglobin a jejich hlavní funkcí je přenos kyslíku do celého organismu, kde ho odevzdá a vezme si oxid uhličitý a odnese ho do plic, kde je následně vydechnutý

Bílé krvinky (leukocyty) - jejich úkolem je boj proti virům, bakteriím i parazitům. Dělí se na:

- Neutrofily (neutrofilní granulocyty)
- Eozinofily (eozinofilní granulocyty)
- Bazofily (bazofilní granulocyty)
- Monocyty
- T-lymfocyty
- B-lymfocyty

Krevní destičky (trombocyty) patří mezi nejmenší krevní elementy, podílí se na zástavě krvácení vytvořením krevní zátky.

Kostní dřev je uložena uvnitř kostí a tvoří se zde všechny typy krvinek. Denně zde z kmenových buněk vznikají nové krevní buňky, které nahrazují staré a poškozené krvinky. [1]

1.2 Typy leukémií

Jako první je třeba určit, jestli se jedná o leukémii akutní nebo chronickou. Dále zda se jedná o myeloidní nebo lymfocytární.

U akutní leukémie dochází k množení zárodečných buněk tzv. blastů v kostní dřeni, které utlačují krvetvorbu ostatních krevních buněk. Nemoc se rozvine během několika měsíců. Pokud nedojde k léčbě, pacient přežívá jen krátce (týdny, měsíce). [1]

U chronické leukémie buňky dozrávají, ale neplní svoji funkci. Pacienti jsou dlouho bez potíží, většinou se nemoc odhalí náhodně, při vyšetření krevního obrazu. Šance na úplné vyléčení je však malá.

Myeloidní a lymfocytární leukémie - druhým znakem je typ kmenové buňky, ze které se leukémie rozvine. Z myeloidní řady jsou to neutrofily a monocyty u lymfocytární jsou to lymfocyty.

- Akutní lymfoblastová leukémie (ALL)
- Akutní myeloidní leukémie (AML)
- Chronická lymfocytární leukémie (CLL)
- Chronická myeloidní leukémie (CML)

1.3 Lymfatický systém a jeho funkce

Je to soustava lymfatických cév a lymfatických orgánů, které procházejí každým místem v lidském těle a jsou spojeny s lymfatickými orgány - brzlík (thymus – pouze v dětství), slezina, kostní dřeň, lymfatické (mízní) uzliny, nosní, krční a patrové mandle. Shluky lymfatické tkáně jsou součástí dýchacího a trávicího ústrojí. V lymfatickém systému jsou nejvíce zastoupeny lymfocyty. V lymfatických cévách proudí míza (lymf) – mléčná nebo bezbarvá tekutina, která vzniká z tkáňového moku v mezibuněčných prostorech. Do cest mízních cév jsou vloženy mízní uzliny. Ty jsou rozmístěny na různých částech těla, viz obrázek č. 1.1.

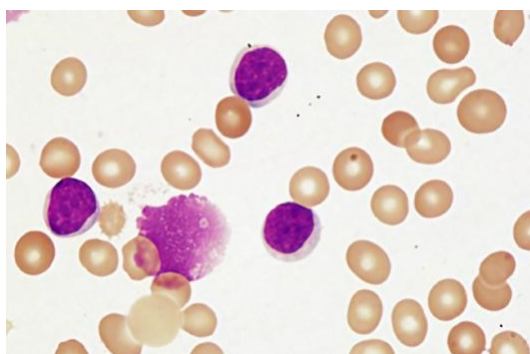
Za běžných okolností nejsou lymfatické uzliny hmatatelné. K jejich zvětšení nejčastěji dochází při nádorových nebo infekčních onemocněních, tedy i u chronické lymfocytární leukémie. Lymfatický systém zajišťuje zpětný transport odpadních látek a vody zpět do krve z tkání a jako součást imunitního systému se podílí na obraně lidského těla proti neinfekčním i infekčním nemocem. [1]



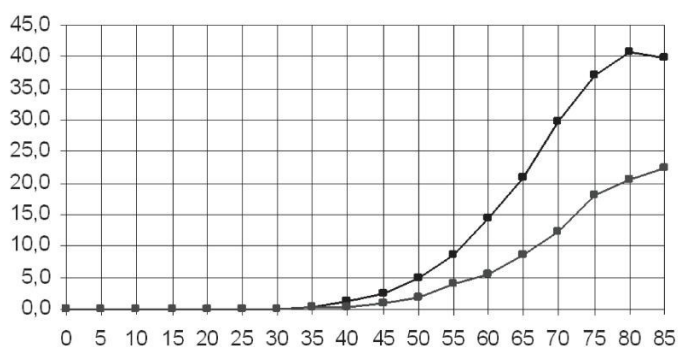
Obrázek č. 1.1: Oblasti mízních (lymfatických) uzlin v lidském těle k věku [1]

1.4 Chronická lymfocytární leukémie (CLL)

Chronická lymfocytární leukémie je lymfoproliferativní nádorové onemocnění, které vzniká z B lymfocytů, které se přeměňují v buňky nádorové s dlouhou životností a také s vyšší schopností se množit než zdravé lymfocyty. Lymfocyty u chronické lymfocytární leukémie jsou vidět na obrázku č. 1.2. Buňky, které jsou nádorově přeměněné, se vymykají kontrole celého organismu, samovolně se množí, neplní funkce, ke kterým byly původně určeny, nepodléhají apoptóze a poškozují organismus. Dochází k hromadění nádorových buněk v organismu (lymfatické uzliny, slezina, játra). Nádorové buňky jsou v kostní dřeni, kde utiskují krvetvorbu, což má za následek, že se netvoří nové krevní buňky. Chronická lymfocytární leukémie postihuje převážně pacienty ve vyšším věku (medián věku při diagnóze je 65-70 let) a medián celkového přežití pacientů je přibližně 10 let. U dětí a osob mladších 25 let nemoc nebyla popsána. CLL má velmi různorodý průběh, někteří pacienti nepotřebují léčbu, nezkracuje jim to život a na druhou stranu k jiným pacientům se chová agresivně a ti přežívají cca 2 - 3 roky od stanovení diagnózy. Touto nemocí výhradně trpí běloši. Jde o nejčastější typ leukémie ve vyspělých zemích (v Evropě a ve Spojených státech amerických). Ženy jsou tímto onemocněním postiženy dvakrát méně než muži. Každým rokem je u nás popsáno asi 6 nových případů této leukémie na 100 000 obyvatel. Výskyt chronické lymfocytární leukémie k věku ukazuje obrázek č. 1.3. [1,2]



Obrázek 1.2: Lymfocyty u chronické lymfocytární leukémie [1]



Obrázek 1.3: Výskyt chronické lymfocytární leukémie v populaci ve vztahu k věku [1]
(na ose y je populace a na ose x věk)

Co je příčinou vzniku chronické lymfocytární leukémie se dosud nezjistilo, ale je zjištěno, že u CLL hraje významnou roli mikroprostředí, v kterém se krvetvorné buňky nacházejí, a změny přímo v krvetvorných buňkách. V nich dochází k mutacím deoxyribonukleové kyseliny (DNA). Tyto mutace se v průběhu času kumulují, až dojde k rozvoji nemoci.

CLL se rozvíjí velmi pomalu, někdy je i desetiletí bezpříznakovým obdobím, patologický krevní obraz je jediný důkaz o nemoci.

Transformace nemoci:

u některých případů se může nemoc transformovat do agresivnější formy, což je Richterův syndrom (nejčastěji difuzní velkobuněčný B lymfom). Dále může přejít do prolymfocytární leukémie (CLL/PLL), nebo vzácně do akutní lymfoblastické leukémie.

1.5 Příznaky

Jako první příznaky může pacient na sobě pozorovat zvětšené lymfatické uzliny či uzlin na krku apod. Projevy mohou být i jiné, jako je bolest a zvětšení břicha, při zvětšené slezině tlak v břiše, pocit plnosti nebo zažívací obtíže při postižení trávicího ústrojí. U pokročilého onemocnění, kvůli kterým přicházejí lidé k lékaři, bývají celkové projevy CLL, do kterých patří: noční pocení, horečka nad 38 °C, celková slabost, větší úbytek hmotnosti (více jak 10 % hmotnosti za půl roku), bledost, hučení v uších, únava, bušení srdce, dušnost nebo bolesti na hrudi. Dále neustálé infekce, poruchy srážení krve apod. [1]

2. Vyšetřovací metody

U chronické lymfocytární leukémie je pro stanovení diagnózy nutné provést řadu vyšetření. Laboratorní vyšetření krve patří mezi základní vyšetření. Při něm je nalezen zvýšený počet bílých krvinek – lymfocytů. Hlavním kritériem diagnózy nemoci je počet lymfocytů charakteristického fenotypu přesahující hodnotu $5 \times 10^9/l$. [1]

Morfologicky ≤ 55 % atypických buněk (např. prolymfocyty) v periferní krvi. Typický imunofenotyp (koexprese znaků CD5+, CD19+, CD23+, CD79b-, atd.).

Lymfom z malých lymfocytů (SLL) s cytopenií z infiltrace kostní dřeně je považován za CLL i bez přítomnosti lymfocytózy. [1]

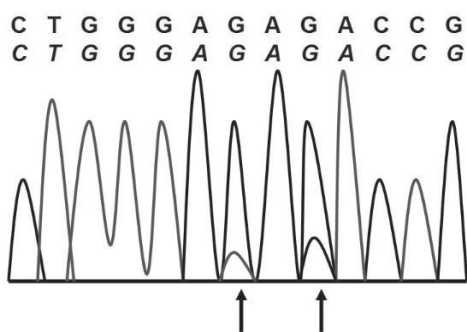
2.1 Základní laboratorní vyšetření pro stanovení diagnózy

Krevní obraz + mikroskopický rozpočet leukocytů, retikulocyty. Minerály, dusíkaté katabolity, jaterní testy, elektroforéza bílkovin, laktát dehydrogenáza (LDH), beta-2 mikroglobulin (B2M) Kvantita sérových imunoglobulinů. Přímý antiglobulinový test (Coombsův test). [1]

Dalším z vyšetření je vyšetření krve na průtokovém cytometru, který detekuje typické CD znaky nádorových buněk na jejich povrchu. Dále molekulárně genetické a cytogenetické vyšetření nádorových buněk, která upřesňují prognózu nemocných. Jak vypadají chromozomy v buňkách chronické lymfocytární leukémie (obrázek číslo 2.1), se zjistí cytogenetickým vyšetřením. Abnormality genů v nádorových buňkách (obrázek 2.2) jsou prokázány molekulárně genetickým vyšetřením. Dalším je podpůrné vyšetření kostní dřeně. V nutném případě se kostní dřeň (dřeňová krev) odebírá z hrudní kosti (sternální punkce; obrázek číslo 2.3) nebo z kosti kyčelní (trepanobiopsie kosti kyčelní; obrázek číslo 2.4). Trepanobiopsií je odebírán nejen vzorek dřeňové krve, ale rovněž malý váleček kostní dřeně široký asi 2 mm a dlouhý 1-2 cm. [1,2]



Obrázek 2.1: Chromozomy v buňkách chronické lymfocytární leukémie zobrazené při cytogenetickém vyšetření. Obrázek ukazuje chybění kousku 17. chromozomu (šipka), které je u chronické lymfocytární leukémie poměrně časté a ovlivňuje prognózu nemoci (fotografie Mgr. Kristina Ďurechová). [1]



Obrázek č. 2.2: Výsledek molekulárně genetické analýzy (sekvenování) ukazující dvě mutace genu *TP53* u chronické lymfocytární leukémie. [1]



Obrázek č. 2.3: Odběr kostní dřeně sternální punkcí [1]



Obrázek č. 2.4: Odběr kostní dřeně trepanobiopsií kosti kyčelní [1]

2.2 Zobrazovací vyšetření při stanovení diagnózy

- Ultrasonografie břicha
- RTG hrudníku
- CT hrudníku a břicha u mladších nemocných
- PET/CT při podezření na Richterovu transformaci

Tímto vyšetřením získané obrazy jsou zpracované počítačem, což umožňuje prostorové zobrazení vnitřních orgánů a tím usnadňuje dokonalejší posouzení stavu jejich postižení. [1]

2.3 Prognostická vyšetření z periferní krve

- Analýza karyotypu buněk nádorové tkáně konvenční cytogenetickou metodou
- Příprava materiálu pro cytogenetické vyšetření

Pro cytogenetické vyšetření používáme buňky kostní dřeně (KD), periferní krve (PK) nebo buňky jiné nádorové tkáně (např.: uzlina).

Odebraný vzorek je uložen do zkumavky s protisrážlivým činidlem (heparinát litný).

KD i PK je kultivovaná in vitro v kultivačním médiu při teplotě cca 37 °C, v 5 % CO₂ po dobu cca 24hod (KD) a 72hod (PK). Kultivace buněk PK probíhá za přítomnosti látek stimulujících lymfocyty k buněčnému dělení EKAMPO MIX (směs IL2 a oligonukleotidů) - pro stimulaci B-lymfocytů. K zastavení dělících se buněk v metafázi přidáváme colcemid.

Cílem kultivace je získat co největší množství dělících se buněk (lymfocytů) pro následující cytogenetické a molekulárně cytogenetické vyšetření. [8]

Příprava a barvení cytogenetických preparátů Giemsovým barvivem

Hodnocení cytogenetických preparátů obarvených Giemsovým barvivem slouží k hodnocení úspěšnosti kultivace biologického materiálu, dále jen BM (množství mitóz získaných po kultivaci BM), nebo k rychlému posouzení počtu chromozómů v mitózách a přítomnosti abnormálních chromozómů. Není však vhodná pro karyotypování vzhledem k nepřítomnosti specifických pruhů na chromozómech.

Barvení na barvicí mřížce:

1. Skla se vzorky se položí na barvicí mřížku, zakápnou se roztokem Giemsa : destilovaná voda v poměru 1 díl barviva Giemsa a 9 dílů destilované vody. Nechá se působit cca 10 minut.
2. Poté se sklo opláchne v kopolně (většinou skleněná hlubší nádoba) s destilovanou vodou a nechá se uschnout při laboratorní teplotě.
3. Následuje mikroskopické hodnocení cytogenetického preparátu při procházejícím světle.

Zpracování biologického materiálu (BM) pro cytogenetické a molekulárně cytogenetické vyšetření u dospělých

Biologický materiál se nejprve zafixuje s metanolem a kyselinou octovou, následně se centrifuguje. Odstraní se supernatant a vzorek se dále inkubuje s KCL. Přidá se roztok Carnoye a opět se centrifuguje. Odstraní se opět supernatant, přidá se vychlazený roztok Carnoye. Následuje opakovaná inkubace, centrifugace až se vytvoří opaleskující mléčně zbarvená suspenze, která je uchována v lednici po dobu několika týdnů. Materiál je dále uchován zamrazením při -20°C. [8]

Příprava a barvení cytogenetických preparátů G-pruhováním a karyotypováním

Metoda barvení cytogenetických preparátů G-pruhováním umožňuje rozlišení jednotlivých chromozómových párů (kondenzované metafázní chromozómy) podle specifických G-pruhů. Hodnocení preparátů lze pomocí mikroskopu nebo počítačových programů. Poté lze sestavit karyotyp a zjistit případné abnormality chromozomů.

1. Do první skleněné kopoliny se připraví barvicí roztok I. smísením:
Na₂HPO₄, kyseliny citronové, Giemsu Karyomax, trypsin
2. Do druhé skleněné kopoliny se připraví barvicí roztok II. smísením:
Leishmanova činidla, Gurruova roztoku

Na podložní sklo se nakape z výšky cca 50 cm 2-3 kapky zpracované buněčné suspenze (v digestoři), nakapané sklo se protáhne nad plamenem a nechá se uschnout při laboratorní teplotě. Takto se připraví sada 5 skel pro každého pacienta a každou kultivaci.

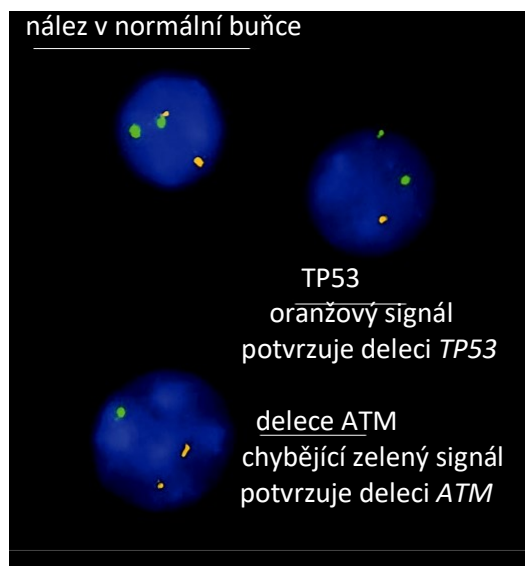
Nejprve se vzorek inkubuje v 1. kopolině, opláchne se roztokem EBSS, následně ve 2. kopolině se opláchuje destilovanou vodou a poté se nechá vzorek uschnout. [8].

Karyotypování a vyhodnocení

Pomocí světelného mikroskopu se vyhledá a vyhodnotí 30 mitóz pro každého pacienta, z 10 z nich se vyhotoví karyotyp za pomoci softwaru Ikaros (MetaSystems).

Fluorescenční in situ hybridizace (FISH)

Chromozomální abnormality v nádorových buňkách lze najít až u 80 % pacientů s CLL. Fluorescenční in situ hybridizace (FISH – delece 13q, 11q, 17p, trisomie 12). Mutační stav genů pro variabilní část těžkého řetězce imunoglobulinu (*IgHV*). Vyšetření mutací genu *TP53*. Vyšetření karyotypu po stimulaci nepříznivé prognostické faktory: delece 17p, nemutovaný *IgVH*, mutace *TP53*, komplexní karyotyp. [2]



Obrázek č. 2.5: Výsledek FISH vyšetření interfázních jader u nemocného s chronickou lymfocytární leukémií. Použita sonda XL *ATM/TP53* (MetaSystems). [2]

Při vyšetření se používá fluorescenčně značené sondy, které se hybridizují s částečně denaturovaným vyšetřovaným chromozomem. Denaturace je proces, kdy se spojí vlákno vyšetřované DNA s druhým vláknem, kterým je značená sonda. Sondy se dělí na centromerické, lokus specifické a telomerické. Obvykle se vyšetřuje metafázní (chromozomy se seřadí do ekvatoriální roviny dělicího vřeténka) chromozomy, které jsou nejlépe kondenzované. Vyšetření lze ovšem provést i na interfázních (tj. období mezi dvěma mitózami (dělením) buňky) jádrech.

Princip

Nejprve musí být vyšetřovaná nukleová kyselina denaturována. Tím se rozruší vodíkové můstky mezi bázemi a dojde k oddělení řetězců. Získáme tak z dvoušroubovice DNA dva jednoduché polynukleotidové řetězce. Podle pravidel komplementarity bází a za reasociačních podmínek je na ně navázán krátký a značený úsek nukleové kyseliny, tzv. sonda. Hybridizační signál se projeví při navázání sondy na vyšetřovanou DNA, který lze pozorovat ve fluorescenčním mikroskopu. [6]

Postup:

Nejprve se připraví sonda, podle chromozómové změny, kterou má metoda FISH prokázat. Nakape se vzorek suspenze na sklo a nechá se volně uschnout. Poté se vloží do roztoku 2x SSC a promyje se v řadě ledových alkoholů a nechá se uschnout na plotýnce při 45°C. Připravená sonda se napipetuje na preparát a přikryje se krycím sklíčkem. Okraje tohoto sklíčka se oblepí rubber-cementem, který zabraňuje vysychání sondy během denaturace a hybridizace v ThermoBrite. Preparát se hybridizuje v hybridizační komůrce při cca 37 °C po dobu cca 24 hod. Odmývá se zbytek nenavázané sondy po hybridizaci v mycích roztocích. Poté se preparát zamontuje a podbarví se interfázní jádra a mitózy na skle. [9]

Hodnocení iFISH:

Preparát se hodnotí ve fluorescenčním mikroskopu za použití příslušných fluorescenčních filtrů. Hodnotí se vždy minimálně 300 jader interfázních buněk, obraz se snímá za pomoci CCD kamery a ukládá se v softwaru ISIS (MetaSystems).

Hodnocení FISH:

Preparát se hodnotí ve fluorescenčním mikroskopu za použití příslušných fluorescenčních filtrů, mitózy vyhledává se buď manuálně, nebo lze využít automatický vyhledávač metafází Metafer4. Reprezentativní počet mitóz je snímán za pomoci CCD kamery a ukládán v systému ISIS (MetaSystems). Nasnímané metafáze se hodnotí za pomoci programu M-FISH (MetaSystems). [9]

Mnohobarevná fluorescenční in situ hybridizace (M-FISH, mBAND)**POSTUP:**

Nakape se vzorek suspenze na sklo a nechá se volně uschnout. Suché sklo se inkubuje v termostatu 20 min při 37 °C a pak se promyje v řadě ledových alkoholů. Poté se inkubuje několikrát a při různých teplotách v roztocích 2x SSC, NaOH, 0,1 SSC a promyje se v řadě ledových alkoholů a nechá se uschnout na plotýnce při 45°C. [10]

Příprava sondy:

Na krycí sklo použijeme cca 8 µl sondy, která se odpipetuje do 0,2 ml eppendorfky.

Denaturace:

Denaturuje se v thermocycleru (program MFISH, mBAND) při teplotě 75°C po dobu 5 minut, pak se sonda ve druhém kroku zchladí na 4°C na 30 s. Nakonec se uchovává 30 minut při 37°C.

Hybridizace:

Napipetuje se denaturovaná sonda na denaturovaný preparát a překryje se krycím sklíčkem. Okraje sklíčka se oblepí rubber-cementem, který zabraňuje vysychání sondy během hybridizace. Hybridizuje se po dobu 24 hod. při cca 37 °C.

Odmývání:

Zbytek nenavázané sondy po ukončení hybridizace se odmyjeme inkubací skla v mycích roztocích.

Po omytí skla se preparát zamontuje a podbarví se interfázní jádra a mitózy na skle. Na sklo se pipetuje cca 10 µl DAPI a překryje se krycím sklíčkem 24x24 mm.

Hodnocení:

Preparát se hodnotí ve fluorescenčním mikroskopu za použití příslušných fluorescenčních filtrů, mitózy se vyhledávají buď manuálně, nebo lze využít automatický vyhledávač metafází Metafer4. Reprezentativní počet mitóz je snímán za pomoci CCD kamery a ukládán v systému

ISIS (Me-taSystems). Nasnímané metafáze se zkaryotypují a hodnotí se za pomoci programu M-FISH, případně mBAND FISH (MetaSystems). [10]

Mutační stav *IgVH*

Princip metody

B-lymfoproliferace vycházejí z jedné transformované buňky, jejíž potomci mají stejnou genetickou výbavu, kterou je možno v oblasti *IgH* identifikovat pomocí polymerázové řetězové reakce (PCR). U chronické lymfocytární leukémie byly definovány dva podtypy onemocnění. První je odvozen z buněk, u kterých se jednalo o setkání s antigenem, bez přítomnosti antigen prezentujících buněk a k somatické mutaci *IgVH* nemuselo dojít (U-*IgVH*, nemutované *IgVH* geny). Druhý podtyp, prognosticky výrazně příznivější, je tvořen lymfocyty, které prošly mutačním procesem v germinálním centru lymfoidních foliklů a na prezentaci antigenu se podílely antigen prezentující buňky (M-*IgVH*, mutované *IgVH* geny). [11]

V případě nalezení monoklonální populace B-lymfocytů pokračujeme dále analýzou somatických hypermutací *IgVH*. Klonální PCR produkty jsou z gelu vyřezány, extrahovány a následně sekvenovány. Mutační status je stanoven srovnáním *IgVH* sekvence od pacienta s germinálními sekvencemi v on-line databázi. Sekvence, které se liší o méně než 2% jsou považovány za nemutované *IgVH* geny. [11]

Postup:

Odebere se venózní krev nebo punktát kostní dřeně do plastové zkumavky s antikoagulačním roztokem K3EDTA a důkladně se promíchá. Po doručení je odběr připraven k izolaci DNA a RNA.

Sekvence získaná z forward primeru je porovnána se sekvencí získanou z reverse primeru a z obou primerů je vytvořena složená sekvence, která je vložena do databáze a porovnávána s germinálními sekvencemi. V databázi je uvedena sekvenční homologie VH subgenu s odpovídajícím germinálním VH subgenem (v procentech). Jedno procento rozdílu mezi analyzovanou sekvencí a sekvencí germinálního VH subgenu koresponduje se třemi nukleotidovými mutacemi v *IgVH* genu. Nemutovaný stav *IgVH* genu je stanoven na základě alespoň 98% sekvenční homologie s odpovídajícím germinálním VH subgenem. Mutační stav *IgVH* je stanoven jako NEMUTOVANÝ, je-li sekvenční homologie s odpovídajícím germinálním VH subgenem (větší než) > 98%, a jako MUTOVANÝ, je-li sekvenční homologie s odpovídajícím germinálním VH subgenem (menší než) < 98%. [11]

Stanovení prognózy

Cytogenetické nálezy (změny chromozomů) a nálezy molekulární se používají v dnešní době ke stanovení prognózy nemoci. Mezi nejčastější chromozomální abnormality patří: delece (chybění) dlouhého raménka chromozomu 13 (del 13q14).

Del 13q bez dalších chromozomálních abnormalit je spojena s dobrou prognózou nemoci. Vyskytuje se monoalelická i bialelická delece nebo také obě. Bialelická delece může zhoršovat

prognózu. Monoalelická delece v méně než 70% může ukazovat na delší dobu do zahájení léčby, u více jak 70% tomu je naopak.

Delece (chybění) dlouhého raménka chromozomu 11 (del 11q22). Tato změna je spojena s horší prognózou a nutností léčby nemoci.

Trizomie chromozomu 12 (jeden nadpočetný chromozom 12). Trizomie 12 je spojena se střední prognózou. Pokud je spojena s jinou chromozomální změnou, může ukazovat na kratší dobu do zahájení léčby.

Delece (chybění) krátkého raménka chromozomu 17 (del 17p13). Na krátkém raménku chromozomu 17 se nachází tumor supresorový (potlačují nádorové bujení) gen *TP53*. Lidé s CLL a del 17p mají vysoké riziko nutnosti léčby a předpoklad ke špatné léčebné odpovědi na většinu chemoterapeutik a některých monoklonálních protilátek. Tito pacienti musí být léčeni odlišným způsobem než zbylí pacienti s CLL. [12]

IgVH gen mutuje fyziologicky v B-lymfocytech a slouží k tvorbě protilátek, aby tyto lymfocyty tvořily protilátky proti konkrétnímu antigenu. Pokud CLL vznikla z B-lymfocytu, který má mutovaný *IgVH* gen, znamená to, že jde o nemoc vniklou ze zralějšího B lymfocytu. Pacienti s mutovaným *IgVH* mají lepší prognózu než nemocní s nemutovaným *IgVH*. [2]

Přestavba 14q32 (*IGH*) gen pro těžký řetězec imunoglobulinů (*IGH*) je další oblastí, která se zapojuje do chromozomálních změn. Přestavby *IGH* jsou spojeny s horší prognózou. Nejčastější translokace jsou t (14,18) s přestavbou *IGH/BCL2* nebo t (14,19) s přestavbou *IGH/BCL3*.

Delece 6q21 postihují dlouhé rameno chromozomu 6 (6q) a patří mezi nejčastější chromozomové aberace u CLL a jsou spojována s horší prognózou. Delece 6q21 se jako primární změna téměř nevyskytuje, spíše jako sekundární změna spojená s progresí onemocnění a kratším OS.

Duplikace genu *MYC* (8q24). Gen *MYC* kóduje onkogenní transkripční faktor, který reguluje až 15% genů v genomu a kontroluje mnoho aspektů buněčného cyklu (DNA replikace, syntéza proteinů, regulace metabolismu, dělení buněk). Duplikace genu je často spojována s nepříznivou prognózou a Richterovou transformací, častěji u mužského pohlaví.

Zmnožení 2p+ je zmnožení krátkého ramene chromozomu 2. Tato změna bývá většinou v pozdějším stádiu onemocnění a bývá spojována s horší prognózou. Samostatně se tato změna vyskytuje vzácně, většinou ji doprovází další chromozomové změny.

Komplexní karyotyp je definován jako přítomnost 3 a více klonálních aberací v karyotypu a je považován za silný prognostický marker rychlé progresy onemocnění a kratším přežitím. Výskyt komplexního karyotypu svědčí o nepřítomnosti mutovaného *IgVH*. [12]

2.4 Klasifikace

K určení stadia (stupně pokročilosti onemocnění) se využívají různá klasifikační schémata – tzv. klinické stážování. Nejčastěji se používá klasifikace podle jejího autora – Kanti Raie. Rozlišuje pět klinických stádií chronické lymfocytární leukemie:

Klasifikace podle Kanti Raie

Stádium 0 - charakterizované pouze zvýšeným počtem lymfocytů v krvi a kostní dřeni

Stádium I - mimo zvýšeného počtu lymfocytů je přítomno i nebolestivé zvětšení lymfatických uzlin

Stádium II – k výše uvedeným nálezům se přidružuje i zvětšení sleziny nebo jater v důsledku jejich postižení nádorovým procesem

Stádium III – dochází k poklesu červeného krevního barviva (anémie; hemoglobin pod 110 g/l)

Stádium IV – dochází k poklesu krevních destiček (trombocytů; krevní destičky pod $100 \times 10^9/l$)

K římským číslicím přidáváme při určení stadia i písmeno A nebo B. A v případě, že nejsou přítomny celkové projevy onemocnění (zvýšená teplota, váhový úbytek, noční pocení) a písmeno B v případě, že tyto příznaky přítomny jsou. Této klasifikaci se podobá systém podle Bineta, jenž je někdy užíván jako alternativa systému Raiova (tabulka 2.1). [1]

Klasifikace podle Jacquese – Louise Bineta

Stádium	Nález v krvi	Postižení uzlin
A	hemoglobin $\geq 100g/l$ a trombocyty $\geq 100 \times 10^9/l$	0-2 postižené oblasti
B	hemoglobin $\geq 100g/l$ a trombocyty $\geq 100 \times 10^9/l$	3 a více oblastí
C	hemoglobin $< 10 g/l$ a/ nebo trombocyty $< 100 \times 10^9/l$	libovolný počet

Tabulka č. 2.1: Klasifikace chronické lymfocytární leukémie podle Bineta. [1]

Velmi důležitý parametr je určení rozsahu nemoci, podle čehož se zahajuje terapie. Léčí se totiž většinou pouze nemocní ve stádiu III nebo IV podle Raie, nebo ve stádiu C podle Bineta, případně nemocné s B symptomy (se zvýšenou teplotou, váhovým úbytkem nebo nočním pocením). [1] Tyto prognostické faktory mají zásadní vliv na délku období bez léčby, období do progresu a celkové přežití, proto byl vytvořen mezinárodní prognostický index, který je využíván.

Proběhla spousta studií na tvorbu prognostických faktorů k vytvoření prognostického skóre, ale zatím se osvědčilo níže uvedené IPI z roku 2016.

2.5 Prognostické modely

Mezinárodní prognostický index (IPI, International Prognostický Index)

Parametr	Bodová hodnota	Riziková skupina	5leté přežití
17p delece/ <i>TP53</i> mutace	4	0–1 Nízké riziko	93,20%
Nemutovaný <i>IgVH</i>	2	2–3 Střední riziko	79,30%
B2M nad 3,5 mg/l	2	4–6 Vysoké riziko	63,30%
Rai I-IV / Binet B a C	1	7–10 Velmi vysoké riziko	63,30%
Věk nad 65 let	1	7–10 Velmi vysoké riziko	23,30%

Tabulka č. 2.2: Mezinárodní prognostický index [2]

Tak, jak je tu už několikrát zmiňováno, u většiny pacientů s chronickou lymfocytární leukémií (CLL) je diagnostikována nemoc v časném stadiu a je léčena aktivním dohledem. Individuální průběh pacientů s CLL v počátečním stadiu je heterogenní a jejich pravděpodobnost nutnosti léčby se při diagnóze těžko předvídá, proto bylo vytvořené prognostické skóre do první léčby (TTFT).

Mezinárodní prognostické skóre pro asymptomatické rané fáze chronické lymfocytární leukémie (IPS-E, International prognostic score for asymptomatic early-stage chronic lymphocytic leukemia) vytvořené na jaře 2020.

Parametr	Bodová hodnota	Riziková skupina	Zahájení léčby - 1 rok	Zahájení léčby - 5 let
<i>IgVH</i> nemutovaný	1	0 Nízké riziko	<1%	8%
Lymfocyty >15 x 10 ⁹ /l	1	1 Střední riziko	3%	28%
Lymfadenopatie	1	2-3 vysoké riziko	14%	61%

Tabulka č. 2.3: Mezinárodní prognostické skóre pro asymptomatické rané fáze chronické lymfocytární leukémie [3]

Externí validace mezinárodního prognostického skóre pro asymptomatické rané fáze chronické lymfocytární leukémie a návrh alternativního skóre (AIPS-E External validation of International Prognostic Score for asymptomatic early stage chronic lymphocytic leukaemia and proposal of an alternative score) vytvořená na podzim 2020.

Tento skórovací systém je alternativou IPS-E, ale AIPS-E kombinuje *IgVH*, fluorescenční in situ hybridizaci a absolutní počet lymfocytů (ALC) > 15 x 10⁹/l. Lymfadenopatie se nahradila fluorescenční in situ hybridizací a tím se eliminuje chyba pozorovatele při detekci mírně zvětšených lymfatických uzlin. AIPS-E lépe odráží biologii onemocnění, konkrétně známou souvislost mezi del (11q) /del (17p) a rozvojem lymfadenopatie. Zároveň zmizí problém detekovat lymfatické uzliny vnitřní (tj. břišní nebo hrudní), které jsou nepřístupné pro klinické vyšetření. [4]

Parametr	Bodová hodnota	Riziková skupina	Zahájení léčby - 5 let
<i>IgVH</i> nemutovaný	1	0 Nízké riziko	14%
Lymfocyty >15 x 10 ⁹ /l	1	1 Střední riziko	41%
FISH del 11q/17p	1	2-3 vysoké riziko	83%

Tabulka č. 2.4: Externí validace mezinárodního prognostického skóre pro asymptomatické rané fáze chronické lymfocytární leukémie a návrh alternativního skóre [4]

2.6 Léčba

Většina pacientů je diagnostikována v časném stádiu onemocnění. Léčba spočívá pouze v pravidelném sledování pacienta. Vzhledem k mimořádné klinické variabilitě nemocných s CLL, a to i v rámci jednoho klinického stádia, je doporučeno u nemocných vyšetření umožňující individuální prognózu zpřesnit. Klinický vývoj těchto pacientů je heterogenní a je obtížné jej předvídat. Někteří pacienti vykazují klidové onemocnění a nikdy nepotřebují léčbu, jiní zase vykazují aktivní onemocnění krátce po stanovení diagnózy a léčbu vyžadují. I když došlo ke zlepšení výsledků u pacientů s CLL, relativní přežití pacientů v rané fázi (tj. Binetova fáze A) se během posledních desetiletí významně nezměnilo. Řízení těchto pacientů představuje velkou výzvu a mělo by být prováděno na základě nejlepších možných prediktivních faktorů a tím určení rizik léčby. [3]

Existuje spousta prognostických modelů, ale nezabývají se pacienty v raném stádiu nemoci, ale celkovým přežitím (OS). Asymptomatictí pacienti nejsou zkoumáni jako samostatná skupina a jejich prediktivní faktory nejsou použity k prognóze vývoje jejich onemocnění. [3]

Teprve na jaře 2020 byl vytvořen prognostický model pro rané stádium nemoci IPS-E, který používá pouze 3 proměnné (*IgVH*, absolutní počet lymfocytů (ALC) $> 15 \times 10^9/l$ a lymfadenopatie) a výsledkem je jejich součet - viz tabulka číslo 3. Informuje nás o pravděpodobnosti, kdy bude potřeba léčby, buď po 1, nebo po 5 letech.

Na podzim 2020 byl vytvořen podobný model AIPS-E, změna byla pouze u parametru lymfadenopatie, která byla nahrazena FISH (del 11q/ del 17p).

Tyto vytvořené modely jsou velkým přínosem pro kliniku, ale jsou stále jen orientační. Chci se pokusit o rozšíření modelu AIPS-E o další faktory jako je věk nad 65 let, komplexní karyotyp (CK), delecí 6q21, trisomii 12, *IgH* a delecí 13q14 a tím zpřesnit fázi do první léčby.

První léčba je u CLL zahájena, jakmile se začnou projevovat klinické příznaky. Léčba je založená zejména na kombinaci chemoterapie a biologické léčby, dále se využívá radioterapie, kortikoterapie, imunoterapie, transplantace kostní dřeně, inhibitory BCR signalizace a imunoterapie monoklonálními protilátkami. Léčba je indikována podle klinického stádia, dynamiky a aktivity onemocnění.

K zahájení léčby se řídíme podle mezinárodně uznávaných kritérií IWCLL 2018. [7]

3. Návrh a tvorba prediktivního modelu pro hodnocení míry rizika léčby u chronické lymfocytární leukémie

V současné době byly vytvořeny 2 prognostické modely do první léčby IPS-E a AIPS-E. Tyto modely určí rizikovou skupinu a procentuální pravděpodobnost do první léčby. Buďto bude potřeba zahájit léčbu po 1, nebo po 5 letech, a to na základě součtu proměnných. Tyto modely jsou orientační, chybí jim však faktory, které by lépe zpřesnily období do první léčby.

3.1 Cíl práce

Hlavním cílem této diplomové práce v praktické části je najít statisticky významné další cytogenetické abnormality, které by ve spojení s AIPE-E zpřesnily predikci do první nasazené léčby. Rozšířit tedy prognostické skóre a vytvořit prediktivní model do první léčby u vybrané skupiny pacientů, kteří mají chronickou lymfocytární leukémii. Jedná se o prognostické faktory jako je věk nad 65 let, komplexní karyotyp (CK), delecí 6q21, trisomii 12, *IgH* a delecí 13q14. V teoretické části, v kapitole 2.5 a 2.6, byly popsány důvody, proč je tento co nejpřesnější model tak přínosný pro pacienty s CLL.

3.2 Soubor dat, statistika a metodika využitá k tvorbě prediktivního modelu

3.2.1. Datová sada

Datová sada obsahuje 100 pacientů s chronickou lymfocytární leukémií, kteří jsou v rané fázi CLL, tzn., že jsou zatím bez léčby, pouze jsou sledováni. Data obsahují věk, pohlaví, komplexní karyotyp (CK), *IgVH*, delecí 6q21, delecí 11q, delecí 17p, trisomii 12, *ACL*, *IgH* a delecí 13q14. Data mi poskytla Hemato-onkologická klinika LF a FN Olomouc. Data jsou anonymní a byla poskytnuta po předchozím informovaném souhlasu pacientů.

Komplexní karyotyp

Je to přítomnost více jak 3 chromosomálních abnormalit v jednom klonu. Statistická analýza CK zpracovaná už v roce 2019 potvrdila negativní vliv jak u TTFT, tak i OS.

Delece 6q21

Naznačuje špatnou prognózu, statistická analýza byla provedená v roce 2017 pouze u OS a ta se prokázala, ale u TTFT analýza provedená nebyla.

Přestavba genu 14q32 (*IgH*)

Spojeno s horší prognózou. Bylo prokázáno ve studiích v roce 2008 kratší přežívání jak u TTFT, tak i u OS.

Delece 13q14

Studie z roku 2017 vykazuje, že pokud se vyskytuje jako samostatná aberace, je spojována s dobrou prognózou a delším přežitím jak u TTFT, tak i u OS. Pokud se nevyskytuje jako samostatná aberace, je tomu naopak.

3.2.2. Statistické zpracování dat

Kaplanův-Meierův odhad funkce přežití

Nejpoužívanějším a nejznámějším neparametrickým odhadem je Kaplanův – Meierův odhad funkce přežití. Kaplanův-Meierův odhad funkce přežití se stal standardem pro hodnocení přežití v klinických studiích. Výpočet je jednoduchý, jde o to, aby byl subjekt v čase t bez sledované události (aby se např. pacient s nádorovým onemocněním dožil času t , nesmí se u něj událost vyskytnout v žádném čase t^* takovém, pro nějž platí, že $t^* < t$). Abychom tedy mohli odhadnout pravděpodobnost, že u daného subjektu se do času t nevyskytne sledovaná událost, musíme odhadnout odpovídající pravděpodobnosti také pro všechny časy t^* , které času t předcházejí. Předpokládejme n různých časů přežití takových, že $t_1 < t_2 < \dots < t_n < t$. Pak pravděpodobnost přežití bez výskytu sledované události až do času t $S(t)$, lze vyjádřit pomocí vztahu:

$$\begin{aligned} S(t) &= P(T > t) = P((T > t_1) \cap (T > t_2) \cap \dots \cap (T > t_{n-1}) \cap (T > t_n)) = \\ &= P(T > t_1)P(T > t_2/T > t_1) \dots P(T > t_{n-1}/T > t_{n-2})P(T > t_n/T > t_{n-1}) \end{aligned} \quad (3.1)$$

Abychom získali odhad $S(t)$, je třeba specifikovat jednotlivé komponenty rovnice. Vzhledem k tomu, že nemáme k dispozici jinou vstupní informaci než pozorované hodnoty, můžeme pravděpodobnost přežití daného času vyjádřit pouze s pomocí údajů o úmrtí v daném čase. Obecně lze tedy psát:

$$\hat{P}(T > t_1/T > t_{i-1}) = \hat{p}_i = \frac{R_i - d_i}{R_i} = 1 - \frac{d_i}{R_i} \quad (3.2)$$

Veličina d_i je počet sledovaných událostí zaznamenaných v čase t_i a R_i je počet subjektů v riziku výskytu sledované události v čase t_i , což je počet subjektů, kteří bez sledované události přestáli čas t_{i-1} . Funkci přežití pak lze odhadnout pomocí vztahu:

$$\hat{S}(t) = \hat{p}_1 \hat{p}_2 \dots \hat{p}_n = \frac{R_1 - d_1}{R_1} \frac{R_2 - d_2}{R_2} \dots \frac{R_{n-1} - d_{n-1}}{R_{n-1}} \frac{R_n - d_n}{R_n} \quad (3.3)$$

Při odhadu pravděpodobností přežití jednotlivých časů t_i je třeba adekvátně zohlednit cenzorování. Cenzorované časy přežití totiž nelze hodnotit stejně jako kompletní pozorování, neboť nepřispívají k d_i , ale zároveň je nelze z hodnocení vyřadit. Kaplanův-Meierův odhad pracuje s cenzorováním tak, že tato pozorování vypadávají ze skupiny subjektů v riziku ihned po zaznamenaném čase cenzorování. Je-li tedy čas t cenzorovaný a platí, že $t_i < t < t_{i+1}$, pak daný subjekt je v čase t_i započítán do skupiny subjektů v riziku $R_{(i)}$, ale v následujícím pozorovaném čase výskytu sledované události t_{i+1} ho již do skupiny v riziku (R_{i+1}) nezahrnuje. Výsledný vzorec pro Kaplanův-Meierův odhad funkce přežití lze tedy jednoduchou úpravou vztahu zapsat jako:

$$\hat{S}(t) = \prod_{t_i \leq t} \hat{p}_i = \prod_{t_i \leq t} \left(1 - \frac{d_i}{R_i}\right) \quad (3.4)$$

Prakticky počítáme výše uvedený součin pouze přes kompletní časy přežití, nicméně teoreticky ho lze definovat přes všechny pozorované časy přežití s tím, že cenzorované časy přežití k odhadu přispívají pouze prostřednictvím R_i , neboť pro cenzorované časy je $d_i = 0$. [13]

Regresivní model

Regresní modelování využívá obecně statistickou metodiku pro hodnocení vztahu mezi vysvětlujícími proměnnými a vysvětlovanou proměnnou. U analýzy přežití je vysvětlovanou proměnnou čas do výskytu sledované události, který je ovšem z důvodu cenzorování problematický, protože v pozorovaných hodnotách se míchají nekompletní a kompletní údaje. Stejně jako pro popisnou analýzu a testování jednoduchých hypotéz o datech přežití tak se musí i pro regresní modelování časů přežití využít nestandardních postupů, které se liší od modelů používaných pro modelování dat bez cenzorovaných hodnot. [13]

Jednou z nejpopulárnějších regresních technik pro výsledky přežití je Coxova regresní analýza proporcionálních rizik. Existuje několik důležitých předpokladů pro vhodné použití Coxova regresního modelu proporcionálních rizik včetně.

1. nezávislost doby přežití mezi odlišnými jednotlivci ve vzorku,
2. multiplikativní vztah mezi prediktory a rizikem (na rozdíl od lineárního, jako to bylo v případě vícenásobné lineární regresní analýzy, podrobněji popsane níže)
3. konstantní poměr rizik v průběhu času.

Coxův regresní model proporcionálních rizik lze napsat následovně:

$$h(t) = h_0(t) \exp(b_1 X_1 + b_2 X_2 + \dots + b_p X_p) \quad (3.5)$$

$h(t)$ je očekávané nebezpečí v čase t , $h_0(t)$ je základní riziko a představuje riziko, když jsou všechny prediktory (nebo nezávislé proměnné) X_1, X_2, X_p rovny nule. Je nutné si všimnout, že predikované riziko (tj. $h(t)$) nebo míra utrpení události zájmu v příštím okamžiku je součinem základního rizika ($h_0(t)$) a exponenciální funkce lineární kombinace prediktorů. Prediktory mají na předpokládané nebezpečí multiplikativní nebo proporcionální účinek.

Je možné vzít do úvahy i jednoduchý model s jedním prediktorem, X_1 . Coxův model proporcionálních rizik je:

$$h(t) = h_0(t) \exp(b_1 X_1) \quad (3.6)$$

Předpokládejme, že bychom chtěli porovnat dva účastníky z hlediska jejich očekávaných rizik, přičemž první má $X_1 = a$ a druhý má $X_1 = b$. Očekávaná rizika jsou $h_a(t) = h_0(t) \exp(b_{1a})$ a $h_b(t) = h_0(t) \exp(b_{1b})$.

Poměr rizik je poměr těchto dvou očekávaných rizik: $h_0(t) \exp(b_{1a}) / h_0(t) \exp(b_{1b}) = \exp(b_{1(ab)})$, který nezávisí na čase, t . Riziko je tedy v průběhu času úměrné.

Někdy je model vyjádřen odlišně, vyjadřuje relativní nebezpečí, což je poměr rizika v čase t k základnímu riziku, k rizikovým faktorům:

$$\frac{h(t)}{h_0(t)} = \exp(b_1 X_1 + b_2 X_2 + \dots + b_p X_p) \quad (3.7)$$

Můžeme použít přirozený logaritmus (\ln) každé strany Coxova regresního modelu proporcionálních rizik, abychom vytvořili následující, které spojuje log relativního rizika s lineární funkcí prediktorů. Za zmínku stojí, že pravá strana rovnice vypadá jako známější lineární kombinace prediktorů nebo rizikových faktorů (jak je vidět v modelu vícenásobné lineární regrese).

$$\ln \left\{ \frac{h(t)}{h_0(t)} \right\} = b_1 X_1 + b_2 X_2 + \dots + b_p X_p \quad (3.8)$$

V praxi spočívá zájem na asociacích mezi každým z rizikových faktorů nebo prediktorů (X_1, X_2, \dots, X_p) a výsledkem. Asociace jsou kvantifikovány koeficienty regresních koeficientů (b_1, b_2, \dots, b_p). Technika odhadu regresních koeficientů v Coxově regresním modelu proporcionálních rizik přesahuje rámec tohoto textu a je popsána v publikacích Cox a Oakes.⁹ Zde se zaměřujeme na tlumočení. Odhadované koeficienty v Coxově regresním modelu proporcionálních rizik, například b_1 , představují změnu v očekávaném logaritmu poměru rizik ve vztahu k jedné jednotkové změně v X_1 , udržující konstantní všechny ostatní prediktory.

Antilog odhadovaného regresního koeficientu, $\exp(b_i)$, vytváří poměr rizik. Pokud je prediktor dichotomický (např. X_1 je indikátorem převládajícího kardiovaskulárního onemocnění nebo mužského pohlaví), pak $\exp(b_1)$ je poměr rizik srovnávající riziko příhody pro účastníky s $X_1 = 1$ (např. převládající kardiovaskulární onemocnění nebo mužské pohlaví) účastníkům s $X_1 = 0$ (např. bez kardiovaskulárních chorob nebo ženského pohlaví). [14]

Pokud je poměr rizik pro prediktor blízký 1, pak tento prediktor nemá vliv na přežití. Pokud je poměr rizik menší než 1, pak je prediktor ochranný (tj. spojený se zlepšeným přežitím) a pokud je poměr rizik větší než 1, pak je prediktor spojen se zvýšeným rizikem (nebo sníženým přežitím). Testy hypotéz se používají k posouzení, zda existují statisticky významné asociace mezi prediktory a dobou do události. Následující příklady ilustrují tyto testy a jejich interpretaci.

Coxův model proporcionálních rizik - se nazývá semiparametrický model, protože neexistují žádné předpoklady o tvaru základní funkce rizika. Existují však i další výše uvedené předpoklady (tj. nezávislost, změny prediktorů způsobují proporcionální změny rizika bez ohledu na čas a lineární souvislost mezi přirozeným logaritmem relativního rizika a prediktory). V analýze přežití se používají další regresní modely, které předpokládají specifické distribuce pro doby přežití, jako jsou exponenciální, Weibull, Gompertzovo a log-normální rozdělení. Například model přežití exponenciální regrese předpokládá, že funkce nebezpečí je konstantní. Jiná rozdělení předpokládají, že se riziko v průběhu času zvyšuje, v průběhu času klesá, nebo se zpočátku zvyšuje a poté klesá. Příklad 5 ilustruje odhad Coxova regresního modelu proporcionálních rizik a diskutuje interpretaci regresních koeficientů. [14]

Hazard ratio (HR) - poměr rizik je statistická hodnota porovnávající riziko výskytu nějaké příhody v závislosti na čase ve dvou různých skupinách, např. v klinické studii se dvěma způsoby léčby při hodnocení doby přežití bez příhody vyjádřené graficky Kaplan-Meierovými křivkami. Počítá se nejen s tím, že u jednotlivých subjektů ve studii příhoda nastane, ale také kdy nastane. HR se udává relativním číslem, pro porovnání léčby A versus B platí, že při HR menším než 1 je A lepší než B. Pokud je např. HR 0,60, často se udává, že léčbou A bylo riziko sníženo o 40 %. [15]

ROC křivka se využívá pro hodnocení a optimalizaci binárního klasifikačního systému (testu), je ukazatelem vztahu mezi senzitivitou a specificitou daného detektoru nebo testu pro všechny přípustné prahové hodnoty. [16]

Plocha pod křivkou (Area Under the Curve) je jedním z kritérií pro výběr vhodného klasifikátoru (modelu) je plocha pod ROC křivkou (AUC), která určuje separaci tříd v daném modelu resp. schopnost separace tříd daným klasifikátorem. AUC má také pravděpodobnostní interpretaci. Při dvou náhodně vybraných vzorcích, z nichž jeden je pozitivní $C_i = 1$ a druhý negativní $C_j = 0$, je AUC pravděpodobnost, že výstup klasifikátoru Y_i u prvního vzorku je většinou výstup Y_j odpovídající druhému vzorku.

$$AUC = P\{Y_i > Y_j\} \quad (3.9)$$

Křivku ROC získáme spojením bodů určených senzitivitou a hodnotou $(1 - \text{specificita})$, které přísluší jednotlivým hodnotám proměnné X .

Podle nejčastěji používaného členění lze test s AUC nad 0,75 již považovat za uspokojivě diskriminující:

AUC: 0,9–1,0 výborně diskriminující

AUC: 0,8–0,9 velmi dobře diskriminující

AUC: 0,7–0,8 dobře diskriminující

AUC: 0,6–0,7 dostatečně diskriminující

AUC: 0,5–0,6 nedostatečně diskriminující. [16]

Youdenova J statistika: často nazývaná také jako Youdenův index, byla použita pro hledání optimální cut – off hodnoty $\max (se + sp)$. Je to jediná statistika, která zachycuje výkon dichotomického diagnostického testu.

Dichotomie je množina rozdělená na dvě části (podmnožin). Jinými slovy, tato dvojice částí musí být společně vyčerpávající: všechno musí patřit jedné nebo druhé části a vzájemně se vylučující: nic nemůže patřit současně do obou částí. [16]

J statistika je: $J = \text{sensitivity} + \text{specificity} - 1$

Přičemž veličinami jsou citlivost a specifita. Vzorec:

$$J = \frac{\text{skutečná pozitivita}}{\text{skutečná pozitivita} + \text{falešná negativita}} + \frac{\text{skutečná negativita}}{\text{skutečná negativita} + \text{falešná pozitivita}} - 1 \quad (3.10)$$

Index dává stejnou váhu falešně pozitivním a falešně negativním hodnotám, takže všechny testy se stejnou hodnotou indexu dávají stejný podíl na celkových nesprávně klasifikovaných výsledcích. Youdenův index se často používá ve spojení s ROC analýzou. Index je definován pro všechny body křivky ROC a maximální hodnotu indexu lze použít jako kritérium pro výběr optimálního mezního bodu, když diagnostický test poskytuje spíše číselný než dichotomický výsledek. [19]

Soubor pacientů:

Celkový počet pacientů:	100
Muži:	62%
Ženy:	38%
Věk > 65 let	69%
ACL > 15 x 10 ⁹ /l	56%
Nemutované IgVH	51%
Delece 11q	24%
Delece 17p	8%
6q21	0,1%
CK	0,2%
IgH	0,1%
13q14 monoal./bialel.	62%
Trisomie 12	0,1%
Léčení pacienti	60%

Ve sledovaném souboru pacientů došlo u některých k zahájení léčby. U těchto pacientů je přesně vypočítaná doba do začátku léčby – tedy konečná data. U ostatních pacientů k léčbě nedošlo, je zde vypočítána jen doba od stanovení diagnózy do data poslední kontroly. Těmto datům se říká cenzorovaná (useknutá) data. Skutečná doba do zahájení léčby by u těchto pacientů byla delší.

Vhodnými metodami pro zpracování cenzorovaných dat jsou Kaplan-Meierova analýza a Coxova regresní analýza. V Kaplan-Meierově analýze se posuzuje, zda mají data vliv na dobu do začátku léčby. V Coxově regresní analýze se může pro každý parametr zjistit, zda je významným prediktorem doby do začátku léčby a toho, zda léčba nastane. Posuzují se kategoriální 0/1 parametry, zde jde tedy o změnu z 0 na 1 (tj. kolikrát se zvýší, resp. sníží riziko TTFT, pokud se 0 změní na 1). Ke statistice HR se dále počítají 95% intervaly spolehlivosti. Je to intervalový odhad této statistiky. Populační hodnota HR leží v tomto intervalu s 95% spolehlivostí. Pro nalezení statisticky významných hodnot (hranic), od jaké hodnoty u daného faktoru je kratší doba do nasazení TTFT, je vhodná Youdenova statistika.

Data jsou statisticky zpracované v programu statistics 22.

Nejprve jsem svá data zpracovala dle výpočtu AIPS-E.

Coxova regresní analýza - odhad statistiky HR pro jednotlivé parametry použité pro výpočet skóre Exp (B) = HR

Variables in the Equation

Parametr	p-value	HR	95% CI	
			Lower	Upper
ACL > 15x10 ⁹ /l	0,118	1,525	0,898	2,588
unmutated <i>IgVH</i>	< 0,0001	6,045	3,054	11,964
del (11q)/ del (17p)	0,001	2,551	1,491	4,362
del (11q)	0,0001	2,948	1,687	5,153
del (17p)	0,621	1,262	0,502	3,171

Tabulka č. 3.1: Určení statistický významných faktorů, které ovlivňují TTFT

Při zpracování dat podle AIPS-E jsem došla k závěru, že:

Parametr ACL > 15x10⁹/l není významným prediktorem TTFT.

Delece (17p) není významným prediktorem TTFT, je-li posuzován samostatně.

Významným prediktorem do nasazení první léčby je unmutated *IgVH* a delece 11q/17p.

		počet	procento
AIPS-E categories	Low-risk	18	20,2%
	Intermediate	33	37,1%
	High-risk	38	42,7%

Tabulka č. 3.2: Kategorizace dle AIPS-E

0 bodů	Low risk
1 bod	Intermediate risk
2-3 body	High risk

Tabulka č. 3.3: Kategorizace rizika podle AIPS-E skóre

Area Under the Curve

Test Result Variable(s)	Area	Std. Error	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
AIPS-E categories	0,764	0,052	< 0,0001	0,662	0,866

Tabulka č. 3.4: C-statistika - AUC u AIPS-E

Původní zpracování AIPS-E je Area 0,74 a 95% interval spolehlivosti je 0,69 – 0,78.[4]

Výsledek původního zpracování Area je srovnatelný s mým výsledkem AIPS-E.

Výpočet 1 a 5 letého TTFT (5 - year cumulative risk of treatment initiation) pro low, intermediate, high risk podle skórovacího systému AIPS-E.

LOW	pst	95%CI	
12měs	11,1%	0%	25,6%
60měs	11,1%	0%	25,6%
INTERMED	pst	95%CI	
12měs	12,1%	9,0%	23,3%
60měs	42,0%	24,4%	59,5%
HIGH	pst	95%CI	
12měs	60,5%	45,0%	76,1%
60měs	85,1%	72,8%	97,4%

s 95% intervaly spolehlivosti

Tabulka č. 3.5: Výpočet 1 a 5 letého TTFT u AIPS-E

V tabulce Life Tables je pro každý specifikovaný čas počet pacientů vystavených riziku TTFT

Number at risk

0	24	48	72	96	120	144	168	192	216	240	months
---	----	----	----	----	-----	-----	-----	-----	-----	-----	--------

Group Low

21	16	15	13	12	9	3	2	1	1	1
----	----	----	----	----	---	---	---	---	---	---

Group

Intermediate

37	26	20	15	13	9	5	2	1	0	0
----	----	----	----	----	---	---	---	---	---	---

Group High

40	12	5	3	2	2	1	0	0	0	0
----	----	---	---	---	---	---	---	---	---	---

Tabulka č. 3.6: Life Tables u AIPS-E

Kaplan – Meierova analýza pro AIPS-E

Odhad průměrné doby do začátku léčby a mediánu doby do začátku léčby pro pacienty s low, intermediate a high rizikem podle prognostického skóre AIPS-E

Means and Medians for Survival Time (months)

AIPSE categories	Mean ^a				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Low risk	206,4	32,9	141,8	270,9	261,7	0,0	.	.
Intermediate	113,6	15,6	83,0	144,2	98,5	23,0	53,5	143,5
High-risk	31,2	8,0	15,5	46,8	11,0	3,0	5,2	16,8
Overall	103,1	14,1	75,4	130,7	50,6	14,6	22,1	79,2

a. Estimation is limited to the largest survival time if it is censored.

Tabulka č. 3.7: Odhad mediánu a průměrné doby do začátku léčby u AIPS-E

Case Processing Summary				
AIPS-E categories	Total N	N of Events	Censored	
			N	Percent
0- low	18	4	14	77,8%
1- intermediate	33	17	16	48,5%
2- high	38	32	6	15,8%
Overall	89	53	36	40,4%

Tabulka č. 3.8: Kategorizace podle míry rizika u AIPS-E

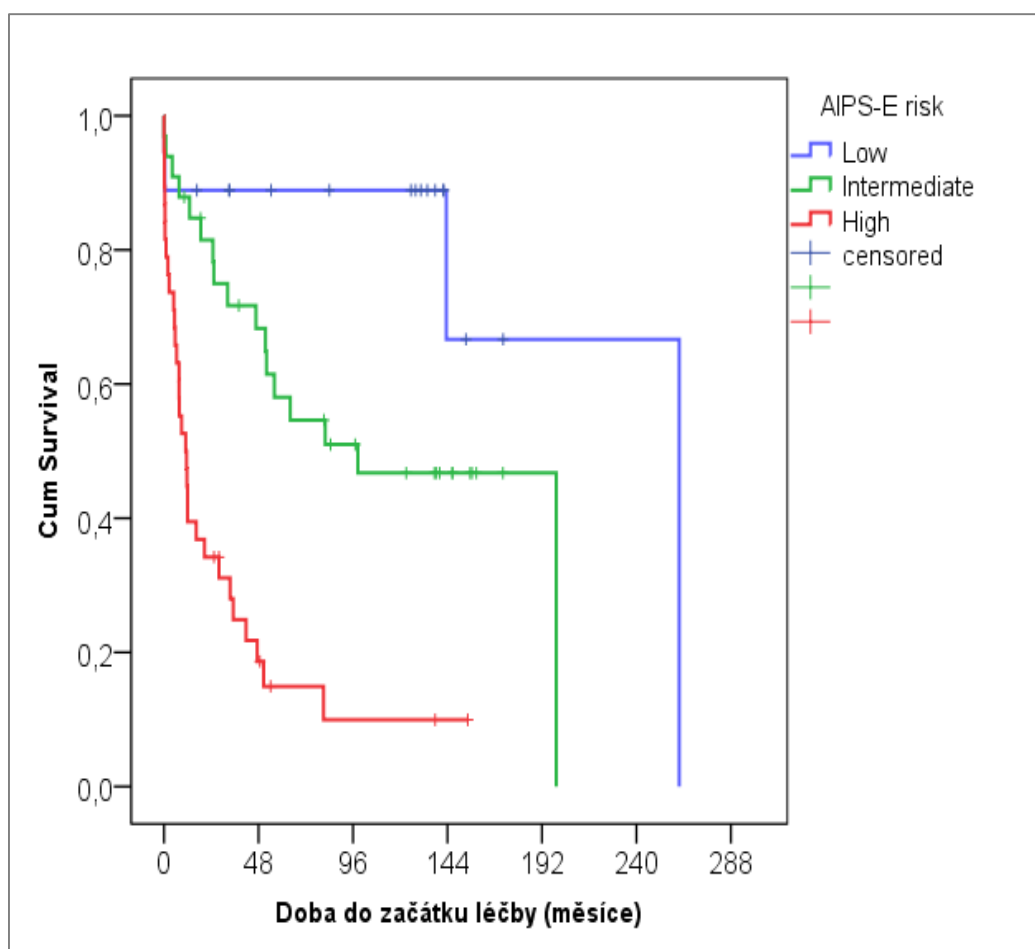
Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	34,375	2	< 0,0001

Test of equality of survival distributions for the different levels of AIPSE categories.

Tabulka č. 3.9: Log Rank test u prognostického skóre AIPS-E

Kaplan – Meierova analýza pro AIPS-E



Obrázek č. 3.1: Kaplan – Meierova křivka doby do začátku léčby u AIPS-E (low – nízké, intermediate – střední a high – vysoké riziko, osa x – doba do začátku léčby, osa y – kumulativní pravděpodobnost)

Z Kaplan – Meierovy křivky je patrné, že pravděpodobnost 0,5 do zahájení první léčby je u vysokého rizika alespoň 11 měsíců, u středního rizika 98,5 měsíců a u nízkého rizika 261,7 měsíců.

Výpočty jsou podobné s původní AIPS-E.

Pětileté kumulativní riziko do zahájení první léčby dle publikace u AIPS-E je 14%, 41%, 83%. [4]

Výpočet AIPS-E u pacientů z FN Olomouc je u pětiletého rizika 11,1%, 42%, 85,1% tabulka č. 3.5.

Nové parametry:

Coxova regresní analýza - odhad statistiky HR pro parametry uvažované pro výpočet prognostického skóre $\text{Exp}(B) = \text{HR}$

U parametrů významných pro predikci TTFT vychází $p < 0,05$.

Variables in the Equation

	p-value	HR	95% CI	
			Lower	Upper
věk nad 65 let	0,187	0,694	0,403	1,195
trisomie 12	0,469	0,731	0,314	1,705
<i>IgH</i>	0,722	1,236	0,385	3,961
delece 6q21	0,004	3,488	1,480	8,218
CK	0,010	2,236	1,217	4,109
Del 13q14 příznivá pouze delece monoalelická, všechny hodnoty %, bez další delece	0,139	0,667	0,390	1,140
Del 13q14 příz. pouze monoalelická <70% a není další delece	0,106	0,632	0,362	1,103
Del 13q14 nepříznivá i bialelelická delece	0,068	0,454	0,194	1,060
Del 13q14 nepříznivá monoalelická delece větší než 70%	0,344	1,361	0,719	2,574

Tabulka č. 3.10: Určení statisticky významných nových faktorů, které ovlivňují TTFT

Uvažované parametry byly: věk nad 65 let, trisomie 12, *IgH*, delece 6q21, CK a delece 13q14. Pokud se delece 13q14 vyskytuje samostatně bez dalších delecí či mutací, ale pouze monoalelická delece a zároveň hodnota menší než 70%, je tato delece pro pacienta příznivá do doby do TTFT. Pokud je hodnota větší než 70%, zhoršuje to prognózu. To se nepotvrdilo. Dále jsem ověřovala, zda u výskytu monoalelické delece a zároveň bialelelické delece a zároveň hodnoty větší než 70% se zhoršuje prognóza do nasazení první léčby. To se taktéž nepotvrdilo.

Statisticky významné parametry pro predikci TTFT jsou pouze parametry delece 6q21 a CK.

Modifikované AIPS-E 1 = AIPS-E + delece 6q21 a CK

Modifikované AIPS-E 2 = AIPS-E + delece 6q21

Modifikované AIPS-E 3 = AIPS-E + CK

	modifik. AIPS-E-1	modifik. AIPS-E-2	modifik. AIPS-E-3
ACL > 15x10 ⁹ /l	1 bod	1 bod	1 bod
unmutated <i>IgVH</i>	1 bod	1 bod	1 bod
del (11q)/ del (17p)	1 bod	1 bod	1 bod
CK	1 bod		1 bod
del 6q21	1 bod	1 bod	

Tabulka č. 3.11: Parametry pro výpočet modifikovaného skóre AIPS-E

	modifik. AIPS-E-1	modifik. AIPS-E-2	modifik. AIPS-E-3
Low risk	0 bodů	0 bodů	0 bodů
Intermediate risk	1 bod	1 bod	1 bod
High risk	2-5 bodů	2-4 body	2-4 body

Tabulka č. 3.12: Kategorizace rizika podle modifikovaného skóre AIPS-E

		počet	procento
modifikované skóre AIPS-E-1, kategorie	Low-risk	16	18,2%
	Intermediate	30	34,1%
	High-risk	42	47,7%
modifikované skóre AIPS-E-2, kategorie	Low-risk	17	19,8%
	Intermediate	30	34,9%
	High-risk	39	45,3%
modifikované skóre AIPS-E-3, kategorie	Low-risk	17	18,7%
	Intermediate	33	36,3%
	High-risk	41	45,1%

Tabulka č. 3.13: Distribuce pacientů podle rizika

Area Under the Curve

Test Result Variable(s)	Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
				Lower Bound	Upper Bound
modifikované skóre AIPS-E-1, kategorie	0,761	0,054	< 0,0001	0,655	0,868
modifikované skóre AIPS-E-2, kategorie	0,767	0,053	< 0,0001	0,663	0,871
modifikované skóre AIPS-E-3, kategorie	0,755	0,055	< 0,0001	0,648	0,863

Tabulka č. 3.14: C-statistika – AUC u modif. AIPS-E

Po přidání k AIPS-E pouze 6q21 vychází AUC lépe než u samotného AIPS-E (AUC 0,764 viz. tab. č. 3.4). Můžeme tedy tvrdit, že přidání delece 6q21 přinese zlepšení predikce. Po přidání CK a nebo CK a delece 6q21 se AUC naopak zhorší, tudíž zhoršuje přesnost prognostického skóre AIPS-E.

Výpočet 1 a 5 letého TTFT (5 - year cumulative risk of treatment initiation) pro low, intermediate, high risk podle skórovacího systémů AIPS-E + 6q21.

LOW	pst	95%CI	
12měs	11,8%	0%	27,1%
60měs	11,8%	0%	27,1%
INTERMED	pst	95%CI	
12měs	13,3%	1,1%	25,5%
60měs	42,7%	24,3%	61,1%
HIGH	pst	95%CI	
12měs	59,0%	43,5%	74,5%
60měs	86,0%	74,0%	98,0%

s 95% intervaly spolehlivosti

Tabulka č. 3.15: Výpočet 1 a 5 letého TTFT u modif. AIPS-E + 6q21

AIPS-E modifikované (AIPS-E + 6q21)

Parametry	Bodová hodnota	Riziková skupina	Zahájení léčby 1 rok	Zahájení léčby 5 let
ACL > 15x10 ⁹ /l	1 bod	nízké riziko (0 bodů)	11,8% (0 - 27,1%)	11,8% (0 - 27,1%)
Unmutated <i>IGHV</i>	1 bod			
Delece 11q/17p	1 bod	střední riziko (1 bod)	13,3% (1,1 - 25,5%)	42,7% (24,3 - 61,1%)
Delece 6q21	1 bod	vysoké riziko (2-4 body)	59,0% (43,5 - 74,5%)	86,0% (74,0 - 98,0%)

s 95% intervaly spolehlivosti

Tabulka č. 3.16: Kategorizace rizika podle bodovacího skóre

V tabulce Life Tables je pro každý specifikovaný čas počet pacientů vystavených riziku

TTFT u AIPS - E + delece 6q21

Number at risk

0	24	48	72	96	120	144	168	192	216	240	months
---	----	----	----	----	-----	-----	-----	-----	-----	-----	--------

Group Low

17	13	12	11	10	7	2	1	1	1	1
----	----	----	----	----	---	---	---	---	---	---

Group

Intermediate

29	22	17	13	12	9	5	2	1	0	0
----	----	----	----	----	---	---	---	---	---	---

Group High

39	13	6	3	2	2	1	0	0	0	0
----	----	---	---	---	---	---	---	---	---	---

Tabulka č 3.17: Life Tables u modif. AIPS-E + 6q21

Means and Medians for Survival Time (months)

Modifikované AIPS-E skóre s delecí 6q21, categories	Mean ^a				Median			
	Estimate	Std. Error	95% Confidence Interval		Estimate	Std. Error	95% Confidence Interval	
			Lower Bound	Upper Bound			Lower Bound	Upper Bound
Low risk	196,2	38,4	120,8	271,5	261,7	0,0	.	.
Intermediate	114,5	16,8	81,5	147,6	98,5	24,6	50,3	146,7
High-risk	31,3	7,7	16,3	46,3	11,5	2,6	6,4	16,5
Overall	100,3	14,2	72,6	128,1	47,3	11,2	25,5	69,2

a. Estimation is limited to the largest survival time if it is censored.

Tabulka č. 3.18: Odhad mediánu a průměrné doby do začátku léčby u modif. AIPS-E + 6q21

Case Processing Summary				
modifikované AIPS-E + del. 6q21	Total N	N of Events	Censored	
			N	Percent
0- low	17	4	13	76,5%
1- intermediate	30	15	15	50,0%
2- high	39	33	6	15,4%
Overall	86	52	34	39,5%

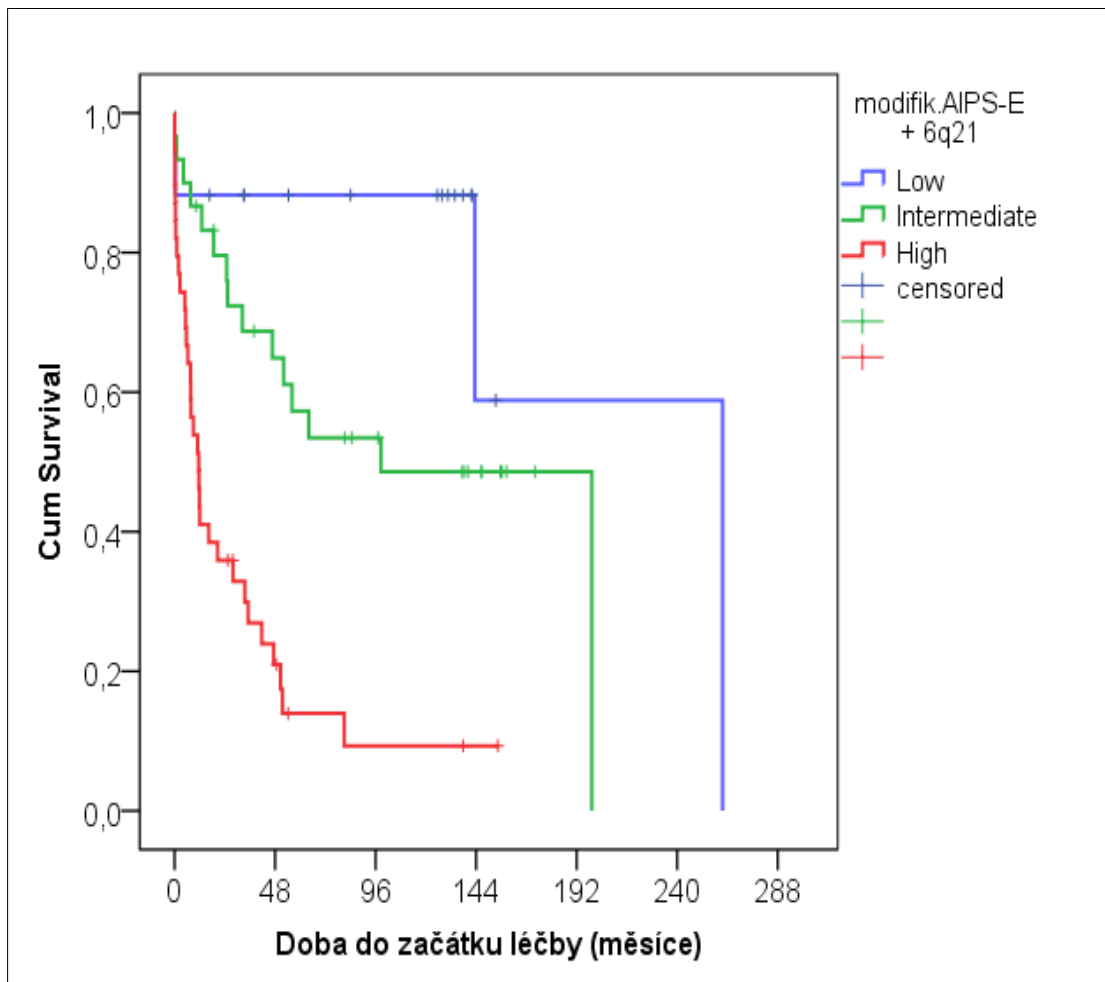
Tabulka č. 3.19: Kategorizace podle míry rizika u mod. AIPS-E + 6q21

Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	30,957	2	< 0,0001

Tabulka č. 3.20: Long Rank test u mod. prognostického skóre AIPS-E + 6q21

Kaplan – Meierova analýza pro AIPS-E + delece 6q21



Obrázek č. 3.2: Kaplan – Meierova křivka doby do začátku léčby u modifikovaného AIPS-E+6q21 (low – nízké, intermediate – střední a high – vysoké riziko, osa x – doba do začátku léčby, osa y – kumulativní pravděpodobnost)

Z Kaplan – Meierovy křivky je patrné, že pravděpodobnost 0,5 do zahájení první léčby je u vysokého rizika alespoň 11,5 měsíců, u středního rizika 98,5 měsíců a u nízkého rizika 261,7 měsíců. Výsledky se velmi podobají hodnotám bez přidání 6q21 k AIPS-E tabulka č. 3.7.

Výsledky jsou srovnatelné.

Pětileté kumulativní riziko do zahájení první léčby dle publikace u AIPS-E je 14%, 41%, 83%. [4] Výpočet AIPS-E u pacientů z FN Olomouc je u pětiletého rizika 11,1%, 42%, 85,1% tabulka č.3.5. Pětileté kum. riziko u mod. AIPS-E + 6q21 je 11,8%, 42,7%, 86% tabulka č. 3.15.

3.2.3. Metodika a tvorba prediktivního modelu

Metodikou k tvorbě prediktivního modelu jsem zvolila fuzzy logiku.

Fuzzy logika

Je matematický podobor logiky, odvozený od fuzzy množin, v němž se logické výroky ohodnocují mírou pravdivosti. Na rozdíl od klasické výrokové logiky nepoužívá pouze dvě logické hodnoty, obvykle zapsané jako 0 a 1. Fuzzy logika využívá všechny hodnoty v rozmezí od 0 po 1 včetně obou hraničních hodnot, kterých je nekonečně mnoho. Pracuje s částečnou pravdou (hodně, málo více apod.). [5]

Fuzzy logika se využívá pro návrhy složitých řídicích systémů. Jejím cílem je vypracovat model vágnosti (příslušnosti). Vágnost tedy značí to, jak moc se proměnná přibližuje pravdivé hodnotě a nabývá hodnot 0 a 1. Pomocí fuzzifikace se dostaneme z reálných hodnot na fuzzy hodnoty. Fuzzifikace je tedy převod reálných proměnných do fuzzy hodnot (např. cena – malá, spíše malá, střední, spíše vyšší, vysoká.

Fuzzy podmnožiny

Musí se definovat vztah mezi fyzikálními proměnnými a fuzzy proměnnými pomocí uspořádaných dvojic: Podmnožina F množiny S je definována jako soubor uspořádaných dvojic, kde první element dvojice je z množiny S a druhý element udává stupeň příslušnosti v intervalu 0,1. Hodnota 0 reprezentuje nepříslušnost, 1 úplnou příslušnost.

Logické operace

Tak jako u dvoustavové logiky tak i u fuzzy logiky, je definován úplný soubor logických funkcí tj. logický součin, logický součet a negace. Ovšem u fuzzy logiky jsou definice obecnější. [5]

Mějme fuzzy proměnné x a y.

Logický součin: $x \text{ and } y = \min(x, y)$; kde min znamená minimum z hodnot uvedených v závorce.

Logický součet: $x \text{ or } y = \max(x, y)$;

Negace: $\text{not } x = 1,0 - x$

Příklad: tedy pro $x = 0,6$ a pro $y = 0,2$ vychází

Logický součin: $x \text{ and } y = \min(x, y) = \min(0,6; 0,2) = 0,2$

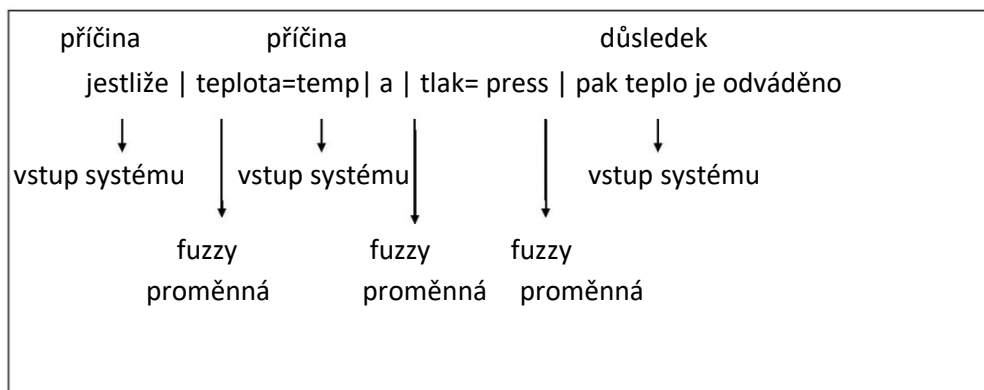
Logický součet: $x \text{ or } y = \max(x, y) = \max(0,6; 0,2) = 0,6$

A pro negace: $\text{not } x = 1,0 - x = 1,0 - 0,6 = 0,4$

Každý vyhodnocovací cyklus jádra generuje systémové vstupní signály, které odpovídají aktuálním podmínkám vstupních fyzikálních veličin. Jádro vyhodnocuje tyto podmínky tak často, jak to vyžaduje aplikace. Proto časování programem je z hlediska výkonu a účinnosti velmi důležité. Každý výpočet se provádí pro konkrétní hodnoty vstupních fyzikálních veličin. [5]

Zpracování má 3 základní fáze – fuzzifikace , vyhodnocení pravidel, defuzzifikace. Defuzzifikace je poslední krok programu, spojuje fuzzy výstupy s výstupem celého systému. K vypočteným fuzzy hodnotám se vyhledají konkrétní hodnoty reálných čísel. [5]

Vyhodnocovací pravidla (příklad)



obrázek č. 3.3: Vyhodnocovací pravidla u fuzzy logiky [5]

Uvedenou strukturu lze číst takto: pokud je teplota vysoká a současně i tlak je vysoký, je nutné odvádět teplo. Proměnné fuzzy jsou označeny jako “temp” a “press”.”[5]

Pro fuzzy model jsem nejprve musela určit hranice hodnot, a to nízká, střední a vysoká hodnota. Hranice jsem určila Youdenovou statistikou, hledání optimální cut – off hodnoty max (se + sp).

Variables in the Equation

Parametr	B	SE	Wald	df	Sig.	Exp (B)	95,0% CI for Exp (B)	
							Lower	Upper
ACL=>22,1 x10 ⁹ /l	0,933	0,264	12,487	1	0,000	2,541	1,515	4,263

Tabulka č. 3.21: Určení hranic ACL podle Youdenovy metody

Pacienti s hodnotou ACL vyšší než 22,1 x 10⁹/l, mají 2,541 krát kratší dobu do TTFT. P- value < 0,001. 95% interval spolehlivosti pro RR je 1,515 – 4,263.

Variables in the Equation

Parametr	B	SE	Wald	df	Sig.	Exp (B)	95,0% CI for Exp (B)	
							Lower	Upper
delece11q>=76,5/delece17p>=44,5	1,301	0,303	18,364	1	0,000	3,671	2,025	6,655

Tabulka č. 3.22: Určení hranic delece 11q a delece 17p podle Youdenovy metody

Pacienti u delece 11q >= 76,5 nebo delece 17p >=44,5, mají 3,671 krát kratší dobu do TTFT. P - value < 0,001. 95% interval spolehlivosti pro RR je 2,025 – 6,655.

Nastavení hranic podle Youdenovy metody u ACL, delece 11q a delece 17p

BODY	HRANICE	delece 11q	delece 17p	ACL > 15x10 ⁹ /l	delece 6q21	unmutated IGHV
0	nízké	0	0	15	0/ne	0/ne
1	střední	76	44	23	1/ano	1/ano
2	vysoké	100	100	300		

Tabulka č. 3.23: Kategorizace nízká, střední a vysoká hodnota

ČETNOSTI	delece 11q	delece 17p	ACL > 15x10 ⁹ /l	delece 6q21	unmutated IGHV
nízké	74	79	44	90	41
střední	10	2	17	6	46
vysoké	13	16	39	0	0
SOUČET	97	97	100	96	87

Tabulka č. 3.24: Kategorizace četnosti podle rozdělení nízká, střední a vysoká hodnota

Četnosti

RIZIKO	NE	1 ROK	5 LET	CELKEM
Nízké 0 (b)	17	3	7	20
1	18	2	10	20
2	16	6	13	22
3	5	10	13	15
4	5	6	11	11
5	4	7	11	11
6	0	1	1	1
Střední 2 (b)	34	8	23	42
Vysoké 3-6 (b)	14	24	36	38

Tabulka č. 3.25 : Kontingenční tabulka rozdělení četnosti do rizikových skupin

	1 rok				5 let			
riziko	bodový odhad %	dolní mez %	horní mez %	Interval %	bodový odhad %	dolní mez %	horní mez %	Interval %
nízké 0 b	15,00	0,00	30,65	(0;31)	35,00	14,10	55,90	(14;56)
střední 2 b	19,05	7,17	30,92	(7;31)	54,76	39,71	69,81	(40;70)
vysoké 3-6 b	63,16	47,82	78,49	(48;78)	94,74	87,64	100	(88;100)

s 95% intervaly spolehlivosti

Tabulka č. 3.26: Bodový odhad podílu pacientů s léčbou započatou do 1 - 5 let

Vzorec pro výpočet intervalu spolehlivosti:

$$(p - z_{1-\alpha/2} \sqrt{p(1-p)/n}; p + z_{1-\alpha/2} \sqrt{p(1-p)/n}) \quad (3.11)$$

Ověření správnosti výpočtů intervalů spolehlivosti podle: Waldův jednovýběrový test.

Ověření přesnosti - úspěšnosti predikce na datech do 1 roku je 75% a do 5 let je 72% u AIPS-E + 6q21

Úspěšnost odhadu predikce	1 rok	5let	CELKEM
CELKEM	75	72	147
CELKEM %	75,00%	72,00%	73,50%

Tabulka č. 3.27: Ověření přesnosti predikce na datech AIPS-E + 6q21

Ověření přesnosti - úspěšnosti predikce na datech do 1 roku je 74% a do 5 let je 72% u AIPS-E Bez 6q21

Úspěšnost odhadu predikce	1 rok	5let	CELKEM
CELKEM	74	72	146
CELKEM %	74,00%	72,00%	73,00%

Tabulka č. 3.28: Ověření přesnosti predikce na datech AIPS-E bez 6q21

Srovnáním tabulek 3.27 a 3.28 je vidět, že přidáním 6q21 k AIPS-E se zlepšil predikce do TTFT.

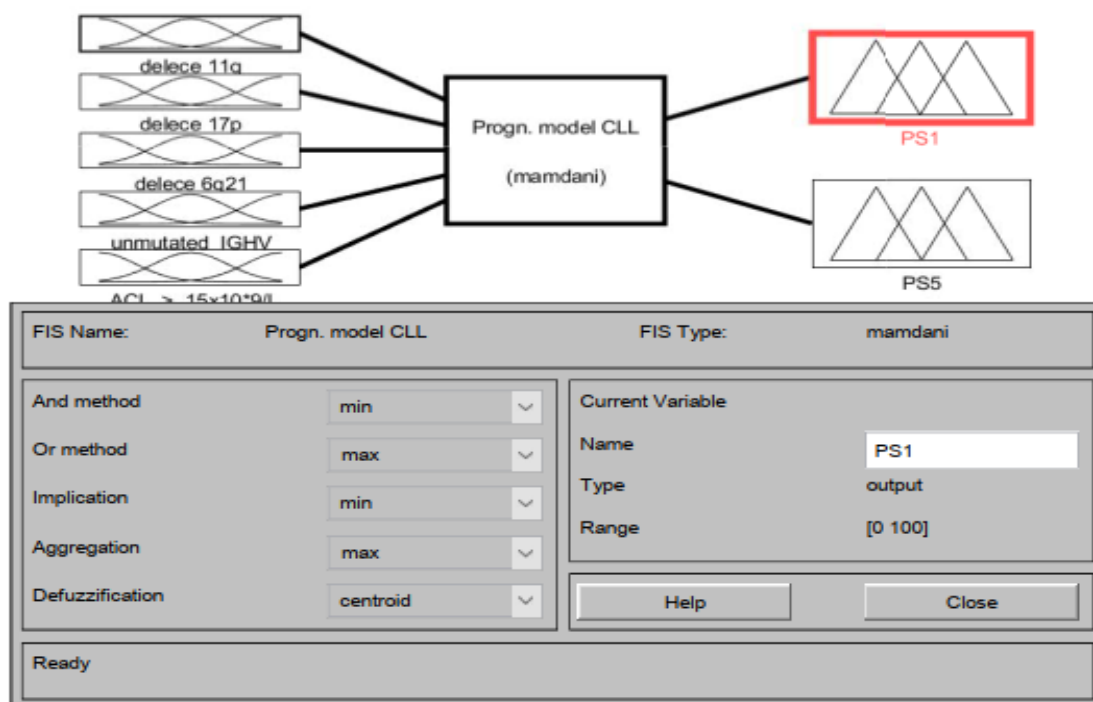
Fuzzy model

delece 11q	delece 17p	delece 6q21	unmutated IGHV	ACL > 15x10 ⁹ /l	Riziko (body)	1 rok	5 let	1 rok	5 let
51	51	0	0	50	vysoké (4)	63,16%	94,74%	(0,48;0,78)	(0,88;1)

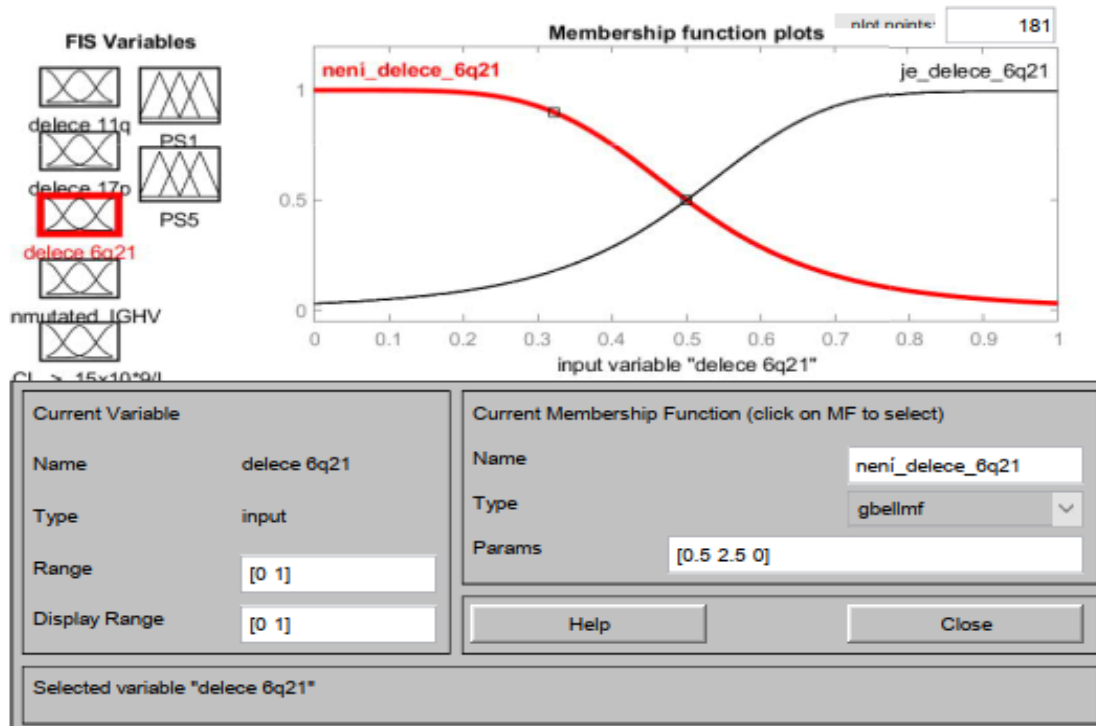
s 95% intervaly spolehlivosti

Tabulka č. 3.29: Výpočet predikce v % pro jednoho pacienta

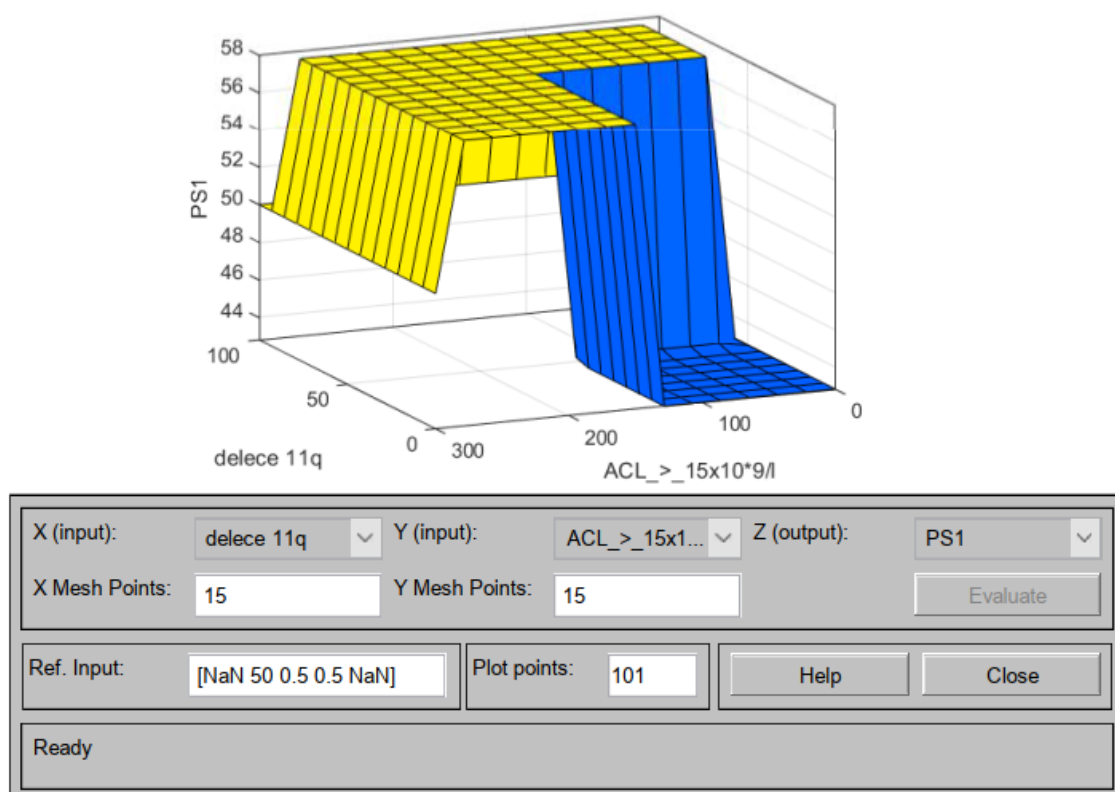
Matlab verze R2021a- Fuzzy Logic Designer



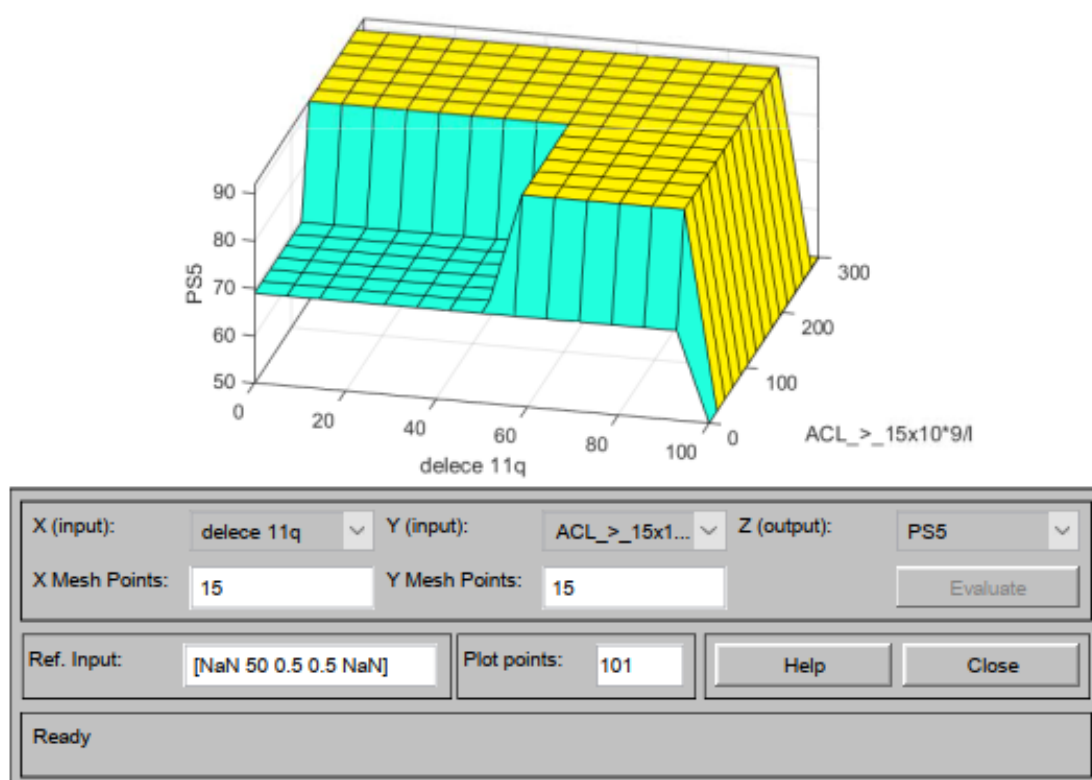
Obrázek č. 3.4: Fuzzy model základní rozdělení



Obrázek č. 3.5: Fuzzy model nastavení členství



Obrázek č. 3.8 Fuzzy model rozložení plochy TTFT do 1 roku



Obrázek č. 3.9 Fuzzy model rozložení plochy TTFT do 5 let

Závěr

V teoretické části jsou uvedeny obecné informace o chronické lymfatické leukémii, vyšetřovací metody, klasifikace, rešerše, metodiky, principy a léčba. Důraz byl kladen zejména na AIPS-E.

Cílem praktické části této práce bylo najít statisticky významné další cytogenetické abnormality, které by ve spojení s AIPS-E přinesly zlepšení predikce doby do nasazení první léčby. Vhodnými metodami pro zpracování cenzurovaných dat jsou Kaplan-Meierova analýza a Coxova regresní analýza. V Kaplan-Meierově analýze se posuzuje, zda mají data vliv na dobu do začátku léčby. V Coxově regresní analýze se může pro každý parametr zjistit, zda je významným prediktorem doby do začátku léčby.

Výsledky této práce ukázaly, že statisticky významné jsou dva faktory a to 6q21 a CK, ale pouze jediný faktor z mých testovacích faktorů ve spojení s AIPS-E, podle C-statistiky, zpřesní predikci doby do nasazení první léčby a to je delece 6q21. Vytvořila jsem dva prognostické modely fuzzy, které vypočítají konkrétní riziko do první léčby v 1. roce nebo v 5 letech pro konkrétního pacienta.

První model je vytvořený v Microsoft Excelu. Tento vytvořený program má svoji výhodu v tom, že po přidání pacienta se celý program na základě četnosti přepočítá, vytvoří se nové % rizika doby do první léčby i intervaly spolehlivosti. Do programu je možné přidávat pacienty se všemi vstupními hodnotami, ale také nemusí. Je možné zadat pouze hodnoty delecí nebo ACL přímo do tabulky, která samostatně počítá z přednastavených hodnot dle už zadaných pacientů a program vypočítá podle přidělených bodů predikci rizika v % do TTFT. V případě zadání všech vstupních hodnot pacienta, je ale výpočet přesnější, jelikož dojde k přepočtu podle četnosti (počet případů, kdy léčba nastala k celkovému počtu pacientů). Nastaví se rizika v procentech a intervaly a pak podle bodové hodnoty je přiděleno riziko v % každému pacientovi. Tento program nepočítá % rizika mezi jednotlivými intervaly, ale vychází z pevně vypočítaných hodnot rizika v % k danému počtu bodů. Samotný výpočet trvá zlomek sekundy. Navržený algoritmus je spustitelný jako soubor s příponou xls.

Druhý program je vytvořený v programu Matlab Fuzzy Logic. Tento program vyniká tím, že počítá % rizika podle přidělených bodů do první léčby i mezi pevně nastavenými intervaly spolehlivosti, což u programu Microsoft Excel není nastavené, dokáže počítat tedy přesněji a výpočet zabere také zlomek sekundy. Zde se pacienti nepřidávají, zadávají se pouze hodnoty delecí nebo ACL a program vypočítá % rizika do TTFT 1 roku i 5 let.

Pokud budu zadávat stejná data do programu Microsoft Excel a programu Matlab Fuzzy Logic, nebudu mít úplně stejné výsledky, jelikož jak jsem už psala program Matlab počítá přesněji, tj. % rizika podle počtu bodů mezi intervaly spolehlivosti. Oba programy pracují spolehlivě, nejsou potřebné žádné speciální nároky na programy ani vybavenost počítače.

Prediktivní model, který představuji v této své diplomové práci, jistě najde své uplatnění v nemocnicích na hemato-onkologických pracovištích.

Seznam použité literatury

- [1] CHRONICKÁ LYMFOCYTÁRNÍ LEUKEMIE, Yvona Brychtová, Anna Panovská, Michael Doubek, Tribun EU, spol. s r.o., 2016, Brno ISBN 978-80-906212-5-1. Dostupné také z <http://www.leukemia-cell.org/res/f/cll.pdf>
- [2] CHRONICKÁ LYMFOCYTÁRNÍ LEUKÉMIE STRUČNÝ PRŮVODCE PRO KAŽDODENNÍ PRAXI, Martin Brejcha, Martin Špaček, Daniel Lysák, Šárka Pospíšilová, Marie Jarošová, Tomáš Papajík, Renata Urbanová, Martin Šimkovič, Lukáš Smolej a Michael Doubek, 2019, ISBN 978-80-270-6266-9 (online pdf) 1. vydání. Dostupné také z https://ccl.cz/sites/default/files/2020-03/CLL_Strucny_pruvodce_pro_kazdodenni_praxi.pdf
- [3] International prognostic score for asymptomatic early-stage chronic lymphocytic leukemia, [online]. 2020 [cit. 2019-9-12]. Dostupné z: <https://www.email.cz/download/k/MONfpw0jFgQSqf9JghVcS8ToqMUvi9RZ4jfcgYUnNiMBGxW2h9tbt-Sbrh3wJ7qpzbPGNTk/International%20prognostic%20score%20for%20asymptomatic%20early%20stage%20CLL%20IPS-E%20Condolucci%20Blood%202020.pdf>
- [4] External validation of International Prognostic Score for asymptomatic early stage chronic lymphocytic leukaemia and proposal of an alternative score, [online]. 2020 [cit. 2019-4-12]. Dostupné z: https://www.email.cz/download/k/Unrwszlb7Vn_22n8XUfgV3DkPPHTelgO3L5EOgzGITDeaebDO8_bLXCnOil8M0a1ub9SQaU/AIPS-E%20Smolej%20BJH%20online%20early.pdf
- [5] Fuzzy logika, [online]. 2011 [cit. 2019-2-12]. Dostupné z: https://moodle.sspbrno.cz/pluginfile.php/5569/mod_resource/content/0/KOLU/Informatika/fuzzy.pdf
- [6] Hybridizace in situ, [online]. 2016 [cit. 2019-2-12]. Dostupné z: https://www.wikiskripta.eu/w/Hybridizace_in_situ
- [7] Doporučení pro diagnostiku a léčbu chronické lymfocytární leukemie (CLL), [online] 2018 [cit. 2019-2-12]. Dostupné z: <https://docplayer.cz/138192965-Doporuceni-pro-diagnostiku-a-lecbu-chronicke-lymfocytarni-leukemie-cll-2018.html>
- [8] SOP FN Olomouc LHOK, Analýza karyotypu buněk nádorové tkáně konvenční cytogenetickou metodou, 2018
- [9] SOP FN Olomouc LHOK, FISH, 2018
- [10] SOP FN Olomouc LHOK, Mnohobarevná fluorescenční in situ hybridizace (M-FISH, mBAND), 2015

- [11] SOP FN Olomouc LHOK, Mutační stav IGHV, 2019
- [12] Chromozomové změny u chronické lymfocytární leukémie, jejich prognostický a prediktivní význam, [online] 2020 [cit. 2021-2-1]. Dostupné z: <https://www.prolekare.cz/casopisy/transfuze-hematologie-dnes/2020-1-14/chromozomove-zmeny-u-chronicke-lymfocytarni-leukemie-jejich-prognosticky-a-prediktivni-vyznam-122015>
- [13] Kaplanův-Meierův odhad funkce přežití, [online] 2020 [cit. 2021-3-21]. Dostupné z: <https://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat--aplikovana-analyza-preziti--neparametricke-odhady--kaplanuv-meieruv-odhad-funkce-preziti>
- [14] Coxova proporcionální analýza rizika regrese, [online] 2016 [cit. 2021-3-21]. Dostupné z: https://sphweb.bumc.bu.edu/otlt/MPHModules/BS/BS704_Survival/BS704_Survival6.html
- [15] Poměr rizik - hazard ratio (HR), [online] 2021 [cit. 2021-3-21]. Dostupné z: <https://www.linkos.cz/slovnicek/risk-ratio-rr-hazard-ratio-hr/>
- [16] Výběr modelu, ROC křivka, [online] 2011 [cit. 2021-3-21]. Dostupné z: https://cw.fel.cvut.cz/old/_media/courses/m33sad/roczadani.pdf
- [17] Softwarová podpora matematických metod v ekonomice a řízení, Petr Sed'a, [online] 2013 [cit. 2021-3-31]. Dostupné z: <https://www.slu.cz/file/cul/f824b294-b7e1-430f-a520-9652f6699b97>
- [18] Interval spolehlivosti pro parametr π binomického rozdělení, [online], [cit. 2021-4-13]. Dostupné z:// <https://portal.matematickabiologie.cz/index.php?pg=aplikovana-analyza-klinickych-a-biologickych-dat--analyza-a-management-dat-pro-zdravotnicke-obory--testovani-hypotez-o-kvalitativnich-promennych--testovani-hypotez-o-podilech--interval-spolehlivosti-pro-parametr-binomickeho-rozdeleni>
- [19] Statistika Youden's J - Youden's J statistic , [online], 2021 [cit. 2021-4-13]. Dostupné z:// https://cs.qaz.wiki/wiki/Youden's_J_statistic